608
17.12.77

THE

QUARTERLY JOURNAL
OF
EXPERIMENTAL PSYCHOLOGY

EDITORIAL BOARD:
O. L. ZANGWILL
D. RUSSELL DAVIS R. C. OLDFIELD

H B. BARLOW
D. E. BROADBENT
A. ELITHORN
R. L. GREGORY
K. R. L. HALL
R. A. HINDE
I. M. L. HUNTER
P. H. R. JAMES
A. R. JONCKHEERE
H. KAY
A. M. UTTLEY
A. J. WATSON
L. WEISKRANTZ
J. W. WHITFIELD

Editorial Assistant: KATHLEEN P. WATTS

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

VOLUME XIV, 1962

PUBLISHED FOR
THE EXPERIMENTAL PSYCHOLOGY SOCIETY
BY
W. HEFFER & SONS LTD. CAMBRIDGE

The Quarterly Journal of Experimental Psychology

Volume XIV

CONTENTS

H. M. B. HURWITZ. Periodicity of Response in Operant Extinction Following Continuous and Fixed Ratio Reinforcement - - - - - - - 1

L. J. HERBERG. Physiological Drives Investigated by Means of Injections into the Cerebral Ventricles of the Rat - - - - - - - 8

J. MACKINTOSH. An Investigation of Reversal Learning in *Octopus Vulgaris* Lamarck - - - - - - - - - - - - 15

A. TREISMAN. Binocular Rivalry and Stereoscopic Depth Perception - - 23

D. R. J. LAMING. A Statistical Test of a Prediction from Information Theory in a Card-Sorting Situation - - - - - - - - - 38

P. C. DODWELL. A Test of Two Theories of Shape Discrimination - - 65

R. CONRAD. Practice, Familiarity and Reading Rate for Words and Nonsense Syllables - - - - - - - - - - - - 71

W. S. ANTHONY, D. H. HOLDING, J. S. LION and W. SLUCKIN. Size-Weight Interaction in Judgements of Compound Stimuli - - - - - 77

P. BURGH and G. C. GRINDLEY. Size of Test Patch and Simultaneous Contrast - - - - - - - - - - - - 89

A. HERON. Immediate Memory in Dialling Performance With and Without Simple Rehearsal - - - - - - - - - - - 94

N. S. SUTHERLAND, A. E. CARR and J. A. MACKINTOSH. Visual Discrimination of Open and Closed Shapes by Rats. I. Training - - - - 129

N. S. SUTHERLAND and A. E. CARR. Visual Discrimination of Open and Closed Shapes by Rats. II. Transfer Tests - - - - - - 140

R. DAVIS. Choice Reaction Times and the Theory of Intermittency in Human Performance - - - - - - - - - - - - - 157

K. GIBBINS and C. I. HOWARTH. The Effect of Intermittent Illumination on the Visual Acuity Threshold - - - - - - - - - 167

J. W. SENDERS and V. L. SENDERS. Woodpile Revisited: A Comment on Gibbins and Howarth's Failure to Repeat Senders' Findings - - - 176

J. Z. YOUNG. Reversal of Learning in Octopus and the Effect of Removal of the Vertical Lobe - - - - - - - - - - - 193

J. Z. YOUNG. Repeated Reversal of Training in Octopus - - - - 206

M. KINSBOURNE and E. K. WARRINGTON. The effect of a Preceding Random Pattern on the Perception of Brief Visual Stimuli - - - - - - 223

M. KINSBOURNE and E. K. WARRINGTON. Further Studies on the Masking of Brief Visual Stimuli by a Random Pattern - - - - - - 235

SHORTER ARTICLES AND NOTES

H. C. A. DALE. On Choosing Between Bets - - - - - - - 49

P. FRAISSE and M. BLANCHETEAU. The Influence of the Number of Alternatives on the Perceptual Recognition Threshold - - - - - - 52

J. R. SYMONS and C. K. MACKAY. An Investigation of Changes in Auditory Sensitivity During the Performance of a Mental Task - - - - 104

R. C. GONZALEZ and M. E. BITTERMAN. A Further Study of Partial Reinforcement in the Turtle - - - - - - - - - - - - - 109

J. BREBNER and I. GORDON. Ensemble Size and Selective Response Times with a Constant Signal Rate - - - - - - - - - - - 113

G. H. MOWBRAY. Some Remarks Concerning Brebner and Gordon's Paper "Ensemble Size and Selective Response Times with a Constant Signal Rate" - - - - - - - - - - - - - - 117

D. W. J. CORCORAN. Noise and Loss of Sleep - - - - - - - 178

R. J. BERGER, P. OLLEY and I. OSWALD. The EEG, Eye-Movements and Dreams of the Blind - - - - - - - - - - - - 183

N. E. WETHERICK. Eliminative and Enumerative Behaviour in a Conceptual Task - - - - - - - - - - - - - - - 246

P. C. WASON. Reply to Wetherick - - - - - - - - - - 250

APPARATUS

J. WILSON. An Apparatus for Recording Figural After-Effects - - - 119

A. H. GREGORY and J. H. DEVINE. An Electronic Stimulus Randomizer - 122

Book Reviews - - - - - - - - - - 58, 125, 187, 251

Publications Received - - - - - - - - - - - - - - 63

Correspondence - - - - - - - - - - - - - - - 192

Proceedings of the Experimental Psychology Society, 1961 - - - - 56

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIV 1962 Part 1

PERIODICITY OF RESPONSE IN OPERANT EXTINCTION FOLLOWING CONTINUOUS AND FIXED RATIO REINFORCEMENT

BY

H. M. B. HURWITZ

From the Department of Psychology, Birkbeck College

The rate of response during an extinction session of two groups of rats, trained under a continuous and a fixed ratio reinforcement schedule respectively, was examined. It was found that both sets of data required a two factor analysis, factor r_A describing the rate of response during response burst periods and factor t_Q the rate of quiescent, non-response periods. Both reinforcement and extinction sessions were analysed with respect to response pattern distributions. It was found that on this variable the two groups differed significantly from each other.

Many attempts have been made to describe mathematically the rate of response during experimental extinction. Perhaps the most favoured description is Ellson's (1939) formulation in which changes of response rate are predicted by a negatively exponential equation.[1]

However, the typical *operant* extinction curve in which cumulative response frequency is plotted as a function of time, shows such gross perturbations as to suggest that multiple factors are required to sustain a quantitative description of the data. Fortunately a technique exists to test whether an array of scores derives from a continuous distribution, and can therefore be described by a monotonic growth function, or whether the scores are sampled from a multimodal distribution. This technique which tests the empirical variance of scores against a theoretical variance has been briefly described in a previous paper (Hurwitz, 1957). There we were able to show that the frequency of response intervals equal to or more than 2 min. is consistently greater from the onset to end of an operant extinction session than predictable by a monotonic function. Such long inter-response intervals occurred with a high probability within the first few minutes of the extinction session and increased in a positively accelerated manner. Support is therefore given to a two factor analysis of extinction data, factor r_A describing the rate of response during response bursts periods and factor t_Q the rate of quiescent, non-response, periods.

The original analysis was undertaken with data drawn from an experiment in which rats were trained under a continuous reinforcement schedule (CRF) where every occurrence of the referent response, lever pressing, provided a reinforcing stimulus, food. It remains to establish whether this two factor approach is equally serviceable in the analysis of extinction data after different reinforcement conditions. The present study compares the extinction data drawn from two groups of animals, one group trained under a CRF schedule, the other group trained under a fixed ratio (FR) schedule.

---

[1] Ellson's equation $r = R\,(1-e^{-bt})$ may be construed as a transformation of an equation to represent changes in inter-response intervals (IRT's),
$IRT = 1/b \log_e(R/R-r)$ where R is the asymptote and r the number of responses.

THE EXPERIMENT

Subjects. Eleven, 3-month-old, hooded rats from the closed colony of animals maintained by the animal laboratory, Department of Psychology, were used. They were allocated to two groups.

Apparatus. The lever pressing apparatus used has been described elsewhere (Hurwitz, 1958). The lever and the food dispensing machine were mounted on the same wall of a 12 in. square box, 5 in. apart. The food dispenser which delivered 0·04 gm. of food into a recessed food tray was placed behind a perspex swinging door which had to be nuzzled open to give access to food. The electrical circuits were arranged so as to deliver food only after the tray door was opened. Each lever press resulted in a buzzer sounding for as long as the lever was held down. The apparatus was enclosed in a sound-proofed chamber.

Records. Lever responses and tray opening responses were recorded on an electronic serial event and time recorder loaned by N. Welford, Cambridge.

Habituation and feeding schedule. The subjects were maintained for three weeks prior to the first experimental session on a feeding schedule which made food available once daily for 2 hr. They were kept on this schedule for the duration of the experiment. These animals, as is customary in our laboratory, had received handling from infancy.

Training procedure. The two groups, designated CRF and FR, were made up of six and five animals respectively. They were given four preliminary training sessions. On each session five reinforcements only were given. On the first session, three pellets were placed into the tray. A further five pellets were then delivered by the experimenter operating an external switch, the presentation of each pellet being preceded by an auditory stimulus (a 0·5-sec. buzz).

On the fifth day, each subject was placed in the apparatus for half an hour. The circuits were programmed to deliver a pellet whenever a lever response was followed by a tray response. If the subject delayed between lever press and tray response for more than 20 sec., the reinforcement programmer was reset automatically. All animals in both groups were first trained on a continuous reinforcement schedule until the criterion of securing a total of 70 reinforcements over two successive sessions was reached. The schedule was then changed for Group FR: reinforcement was made contingent on emitting three lever responses. This arrangement was maintained for five half-hour sessions. Group CRF, on the other hand, were given a further five half-hour sessions under a continuous reinforcement schedule.

The final phase involved experimental extinction and lasted for two half-hour sessions only. During this period the lever continued to operate the buzzer but the food magazine was emptied of pellets.

RESULTS

Rate of response during reinforcement sessions

The rate of lever and tray responding during the final two reinforcement sessions for Group CRF and Group FR as also the number of reinforcements per session are given in Table I. The rate of lever response is significantly different for the two Groups, Group $\mathrm{CRF_{mean}} = 64{\cdot}9$ vs. Group $\mathrm{FR_{mean}} = 127{\cdot}4$. These results are in line with general findings on the effect of fixed ratio schedules on response rate (see Boren, 1953). The difference in rate of tray responding is not significant, however, Group $\mathrm{CRF_{mean}} = 128{\cdot}6$ vs. Group $\mathrm{FR_{mean}} = 141{\cdot}2$. A Mann–Whitney test for small samples (Siegel, 1956) performed on the number of reinforcements obtained, shows Group CRF to have secured more reinforcements per session than Group FR (Group $\mathrm{CRF_{mean}} = 55{\cdot}0$ vs. Group $\mathrm{FR_{mean}} = 40{\cdot}8$, $p = 0{\cdot}01$).

TABLE I

FREQUENCY OF REINFORCEMENT, LEVER AND TRAY RESPONSES ON THE FINAL TWO TRAINING SESSIONS AND FREQUENCY OF LEVER AND TRAY RESPONSES ON THE TWO EXTINCTION SESSIONS

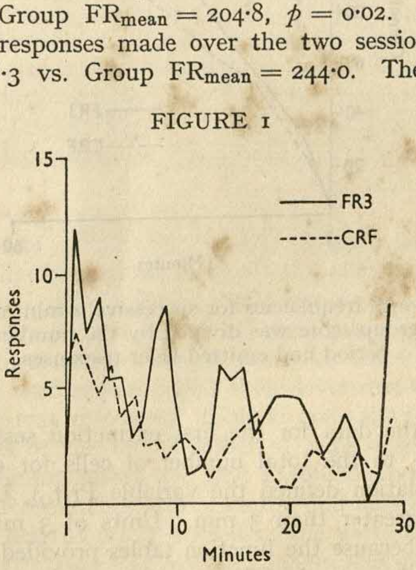
CRF Training										
Semi-final training trial				*Final training trial*			*Extinction I*		*Extinction II*	
S	*L*	*T*	*N*	*L*	*T*	*N*	*L*	*T*	*L*	*T*
1	70	147	47	60	148	51	119	196	72	127
2	60	128	59	83	126	78	81	129	41	129
3	51	117	42	57	160	40	84	112	53	79
4	76	124	68	64	105	50	84	124	44	94
5	67	126	61	78	125	68	41	111	30	89
6	78	152	63	35	86	34	109	119	54	128
FR Training										
1	140	183	46	147	158	46	171	229	113	129
2	112	144	43	156	171	51	92	150	53	50
3	127	162	40	143	189	47	115	133	47	54
4	110	94	36	54	109	18	120	117	86	128
5	125	92	34	160	110	47	127	129	106	100

CRF—continuous reinforcement; FR—fixed ratio reinforcement; L—lever responses; T—tray responses; N—number of reinforcements; S—subjects.

Rate of response during extinction sessions

The difference in the number of lever responses over the two experimental extinction sessions between Group CRF and Group FR is significant, Group $\text{CRF}_{\text{mean}} = 135{\cdot}3$ vs. Group $\text{FR}_{\text{mean}} = 204{\cdot}8$, $p = 0{\cdot}02$. A similar analysis for total number of tray responses made over the two sessions was not significant, Group $\text{CRF}_{\text{mean}} = 248{\cdot}3$ vs. Group $\text{FR}_{\text{mean}} = 244{\cdot}0$. The results for the lever

FIGURE 1

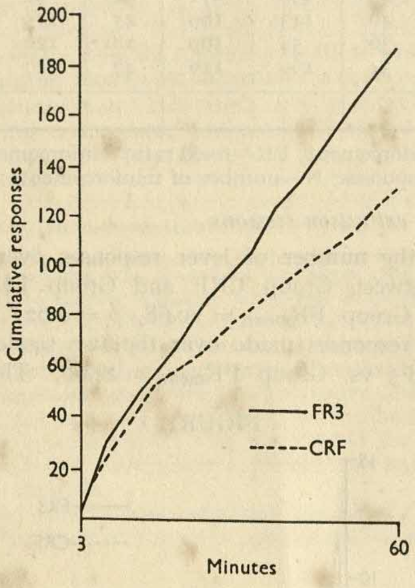


Mean response frequency/minute during the first experimental extinction session for Group CRF, continuous reinforcement trained, and Group FR, fixed ratio reinforcement trained.

response rates are in line with related studies on the effects of schedules of reinforcement on extinction scores (Jenkins and Stanley, 1950; Keller and Schoenfeld, 1950; Ferster and Skinner, 1957). Figure 1 represents the response frequency for both groups as a function of time in the first extinction sessions. The graph shows that by and large the rate of response is consistently higher following fixed ratio reinforcement than following continuous reinforcement training.

A further analysis of lever scores identical to the procedure described in the earlier study (Hurwitz, 1957) was undertaken. The two extinction sessions were treated together for purposes of this analysis and were divided into 30 two-min. periods. The resultant array of cells was examined and all cell entries of 0 and 1 were regarded as indicative of quiescent periods, t_Q, and all entries of above 2 as active periods, t_A. Figure 2 shows the cumulative response frequencies for successive t_A's plotted as a function of time. The curves are slightly decelerated but diverge from each other, thus indicating (*a*) a gradual decrease in response rate during the sessions for both groups, (*b*) an overall slower rate of response for Group CRF.

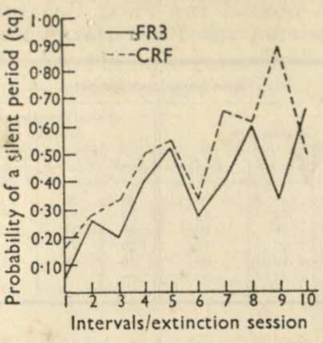
FIGURE 2



Cumulative (mean) response frequencies for successive 2-min. active periods. For each point plotted the total group score was divided by the number of subjects who in that period had emitted lever responses.

Figure 3 presents the data for the first extinction session only in a different form. The ratio of t_Q to the total number of cells for each 3-min. period was computed. This calculation defined the variable $Pr(t_Q)$, i.e. the probability of a non-response interval greater than 3 min. Units of 3 min. were chosen, rather than 2 min. as above, because the function tables provided by Bush and Mosteller (1955) for comparing theoretical and empirical variances contain 10 steps only. The graph shows that $Pr(t_Q)$ increased as a function of time elapsed and that the rate of increase of $Pr(t_Q)$ is slightly greater for Group CRF than for Group FR.

FIGURE 3



The probability of a silent, quiescent period, $Pr(t_Q)$ for Group CRF and Group FR plotted as a function of successive periods during experimental session 1. Each division of the abscissa represents a 3-min. period.

Two factor analysis

The rate of response curves shown in Figure 1 (extinction session 1) indicate considerable changes in rate over successive periods. The present analysis sought to establish whether the general change in rate could be predicted by a monotonic growth function, or whether the present data, particularly the rate for subjects trained under a fixed ratio reinforcement schedule, exhibits the same bi-modal trends previously observed for continuous reinforcement trained subjects.

Each subject's extinction scores (extinction 1 scores only were used) were grouped into 3-min. intervals. By using a maximum likelihood procedure, the variance of the resultant array of scores was compared to the variance of a theoretical array of scores which had been calculated on the assumption that the rate of change in scores declines in a monotonic manner. This technique makes use of a function defined by Bush and Mosteller (1955), namely $g_r(d)$. Thus, given the initial rate of response for each subject, the computations permit a comparison between an empirical and a theoretical distribution. Calculations showed that the variance for both Group CRF and Group FR exceeded the limits allowed, i.e. that a monotonic function could not be fitted to express changes in rate of response.

Run length distributions

In previous experiments using a two response situation like that described here, the run length measure has been found to be a useful additional dependent variable (Fechner, 1958; Hurwitz, 1958). Where reinforcement is contingent on emitting two classes of response, say R_A and R_B, a run length may be defined as the number of successive responses of class R_A before a change to the second class, R_B, occurs. The fixed ratio schedule used required that three lever responses be emitted before reinforcement followed a tray response. Table II gives the run length distributions expressed as percentages of the total number of runs observed, for each subject during both the final reinforcement and first extinction session. The run length distribution for Group CRF matches, within limits, that already observed in a subsequent study where run lengths rarely exceed 1: the pattern appears to be one of strict alternation. In extinction this pattern changes and gives way to longer and longer runs of both lever and tray responses (see Millenson and Hurwitz, 1961). Group FR, by contrast, shows that in many cases run lengths greater than 1 are observed during the reinforcement session (Fechner, 1958). Subsequent experiments

TABLE II

RUN LENGTH DISTRIBUTIONS DURING THE FINAL CONTINUOUS AND FIXED RATIO REINFORCEMENT AND DURING THE FIRST HALF-HOUR EXTINCTION SESSIONS

Run length distributions													
Final reinforced Session			*CRF S*	*Extinction*			*Final reinforced Session*			*FR S*	*Extinction*		
L→T	*LL→T*	*>LLL→T*		*L→T*	*LL→T*	*>LLL→T*	*L→T*	*LL→T*	*>LLL→T*		*L→T*	*LL→T*	*>LLL→T*
89	08	02	1	70	25	05	35	53	12	1	62	27	11
94	05	00	2	62	24	13	73	22	05	2	33	41	26
94	06	00	3	54	26	19	80	17	03	3	17	25	58
96	03	00	4	43	42	16	89	11	00	4	10	50	40
92	06	01	5	81	12	06	10	69	21	5	03	43	34
96	04	00	6	71	22	07	—	—	—	—	—	—	—

(Hurwitz, Brener and Jones, 1962) in which training on different fixed ratio schedules was carried on for much longer periods, have shown that run length distributions peak at the values specified by the reinforcement schedules. During extinction, run length of three (or more) successive responses on the lever significantly exceed the relative frequency of such lengths observed during reinforcement. A comparison of run length distributions under extinction between Group CRF and Group FR showed that Group FR makes significantly more long runs than Group CRF (the relative frequencies for category LLL → T in Table II were ranked and tested by a Mann–Whitney test; $p < 0{\cdot}01$).

DISCUSSION

The aim of the present study was to investigate whether the conclusion reached on a previous occasion (Hurwitz, 1957) that the temporal distribution of extinction scores following continuous reinforcement training (CRF) is most appropriately described by two factors, one referring to the rate of response during response burst periods, the other describing the incidence of quiescent periods, holds also for fixed ratio (FR) trained subjects. The fixed ratio schedule used required three lever presses to ensure reinforcement. The results show that despite the overall higher extinction response rate following FR training, the temporal distribution of responses is clearly bi-modal. It was also shown that such long breaks in response bursts occurred more slowly for FR subjects than CRF ones.

In interpreting these results reference must be made to the finding by Boren (1953) in his extensive study of behaviour under different fixed ratio schedules, that the rate of response during reinforcement sessions is discontinuous. It appears that there are breaks after reinforcement, the size of the break increasing with the size of the ratio employed. Although Boren, too, observed alternations between periods of rapid responding and slow responding during extinction sessions, no attempt was made to establish whether the break periods were in any way correlated with the pause after reinforcement periods during training sessions.

Now, the explanation of the break after reinforcement phenomenon may well be along the following lines: the fixed ratio schedule is an *indirect* method of spacing reinforcements. A more satisfactory arrangement, where intervals between successive reinforcements would be under experimental control, is the fixed interval type schedule. Here it is observed that pause length is a function of the size of the interval. A related finding (e.g. Millenson, 1959) comes from studies where the intervals between successive reinforcements are controlled as also the portion of time within each interval during which the response—which theoretically could be emitted at a constant rate throughout the interval—eventuates in a reinforcement.

However, none of these studies attempts to correlate the inter-response intervals observed during reinforcement sessions with their distributions during experimental extinction. The present result suggests a negative correlation, but clearly more experimental evidence is needed before one can commit oneself to this conclusion. It is also clear that such a study would require different kinds of data from that reported here: it would necessitate recording successive inter-response times which could yield proper distributions (e.g. Millenson and Hurwitz, 1961).

Thus the more general question is raised of the relation between response measures obtained under reinforcement and extinction. One promising lead comes from an examination of run length distributions. Here a positive relation between the effects of reinforcement and extinction are seen since subjects trained by an FR schedule tended to emit significantly longer runs during both phases of the experiment than CRF trained animals.

The experiments were supported by a grant in aid from the Central Research Fund, University of London, 1956.

REFERENCES

BOREN, J. J. (1953). Response rate and resistance to extinction as functions of the fixed ratio. *Ph.D. dissertation, Columbia University.*
BUSH, R. R., and MOSTELLER, F. (1955). *Stochastic Models for Learning.* New York.
ELLSON, D. G. (1939). The concept of reflex reserve. *Psychol. Rev.*, **46**, 566–75.
FECHNER, F. (1958). Probability relations within response sequences maintained under ratio reinforcement. *J. exp. anal. Behav.*, **1**, 109–21.
FERSTER, C., and SKINNER, B. F. (1957). *Schedules of Reinforcement.* New York.
HURWITZ, H. M. B. (1957). Periodicity of response in operant extinction. *Quart. J. exp. Psychol.*, **9**, 177–84.
HURWITZ, H. M. B. (1958). A source of error in estimating the number of reinforcements in a lever pressing apparatus. *J. exp. anal. Behav.*, **1**, 149–52.
HURWITZ, H. M. B., BRENER, J., and JONES, B. (1962). An investigation of two measures of behaviour under FR schedules of reinforcement (in preparation, see also *Bull. Brit. Psychol. Soc.*, 1961, p. 2A).
JENKINS, W. O., and STANLEY, W. C. (1950). Partial reinforcement: a review and critique. *Psychol. Bull.*, **47**, 194–234.
KELLER, F. S., and SCHOENFELD, W. (1950). *Principles of Psychology.* New York.
MILLENSON, J. R. (1959). Some behavioural effects of a two valued temporally defined reinforcement schedule. *J. exp. anal. Behav.*, **3**, 191–202.
MILLENSON, J. R., and HURWITZ, H. M. B. (1961). Some temporal and sequential properties of behaviour during conditioning and extinction. *J. exp. anal. Behav.*, **4**, 97–106.
SIEGEL, S. (1956). *Nonparametric Statistics.* London.

PHYSIOLOGICAL DRIVES INVESTIGATED BY MEANS OF INJECTIONS INTO THE CEREBRAL VENTRICLES OF THE RAT

BY

L. J. HERBERG*

From the Institute of Experimental Psychology, Oxford University

A technique is described for permanently implanting a cannula in the lateral ventricle of the rat, and the effects on certain physiological drives of various substances injected by this route were investigated.

Food intake was significantly reduced by intraventricular glucose, while intraventricular fructose, insulin, glucagon and a lipid-mobilizing agent were without effect. Drinking was elicited by hypertonic saline and suppressed by hypotonic solutions. The intake of salt was not effected by intraventricular low-sodium solutions. Intraventricular procaine, but not alcohol, induced an akinetic state. Intraventricular glucose failed to relieve hypoglycaemic coma.

The significance of some of these findings is discussed in relation to the hypothalamic regulation of hunger.

INTRODUCTION

In the past few years lesion studies and electrical stimulation of the diencephalon have focussed attention on several hypothalamic areas concerned in the arousal and satiation of physiological drives. Another line of investigation has been suggested by the technique developed by Feldberg and Sherwood (1954*a*) who studied the effects of drugs injected via a cannula permanently implanted in the lateral ventricle of the unanaesthetized cat. It was found, for example, that intraventricular barbiturates caused voracious over-eating (Feldberg, 1957), probably by inhibiting the hypothalamic ventromedial nucleus. In another study, Andersson (1953) induced drinking in goats by administering acute intraventricular injections of hypertonic saline which presumably stimulated hypothalamic osmoreceptors. Evidence will be presented that several other substances injected by this route act primarily on the hypothalamus, and the present experiments, using a similar technique, were designed to investigate the biochemical regulation of hunger, satiation, thirst, salt-hunger and certain aspects of general drive level.

Rats have several advantages over cats for behavioural studies, being freely available and more readily trained, and their behavioural characteristics are better known. A simple technique permitting repeated intraventricular injection of the rat will be described in detail.

METHOD

The intraventricular cannulas were prepared from size 21 s.w.g. stainless steel hypodermic needles ground down to a length of 8·0 mm., including 1·0 mm. terminal bevel. They were mounted in 3·0 mm. diameter perspex disks punched out of 1 mm. sheets, so that the cannula point extended 4·0 mm. below the disk. A size 9 rustless sewing-needle clamped into the electrode holder served as a trocar for mounting the cannula on a Krieg stereotaxic instrument.

The implantation which was performed under ether anaesthesia and antibiotic cover, involved making a 1·5 cm. left parasagittal scalp incision and drilling a small burr-hole through the exposed skull at a point 1·5 mm. posterior and 2·7 mm. to the right of bregma. The sewing-needle holding the cannula on its upper length was lowered through the burr-hole for a distance of 2 to 3 mm., and the cannula was then gently pushed down the length of the needle until the perspex mount was in contact with the skull. The perspex and surrounding skull surface were then allowed to dry and covered with dental cement. Three 16BA cheese-head bolts were screwed into the skull to provide additional purchase for the dental cement. When the cement had hardened the trocar was withdrawn and the scalp incision sutured, the cannula protruding through an appropriately placed stab-wound.

---

* Leverhulme Mental Health Research Fund Fellow.

A 2-week period was allowed for recovery before administering the first intraventricular injections, which were given with a Wellcome "Agla" microsyringe coupled to a synchronous electric motor turning at 0·5 revolutions per minute. This gave an injection rate of 0·005 ml./min., which could be maintained for long periods without giving rise to signs of increased intracranial pressure. In practice, a standard injection of 0·2 ml./kgm. body weight was always given, except where otherwise stated.

The syringe was connected to the unrestrained rat by a 2 ft. length of sterilised nylon tubing, the end of which had been drawn and tapered over low heat so as to fit snugly and reach down the whole length of the cannula lumen. When inserting the end of the tubing into the bore of the cannula care was necessary to prevent the sterile tip from coming into contact with the animal's scalp. This was best accomplished by loosely supporting the rat without forcible restraint on the floor of the cage with one hand and inserting the tubing with the other. After a few days of handling, the rats were almost always sufficiently docile to allow this procedure to be completed in a matter of seconds. Occasionally some animals were distracted by the inserted tubing dangling above their heads, but this difficulty was readily overcome by leaving a coiled length of discarded tubing in the cage overnight. The animals habituated in this way did not interfere later with the injection tubing proper.

Between injections, the cannulas were sealed off with tapered plastic bungs moistened with antiseptic. These were prepared from size 000 nylon rod cut to an appropriate length so that when inserted they reached down the full length of the cannula, thus maintaining a patent lumen. They were made to project 2 or 3 mm. above the upper end of the cannula and could be readily extracted with the aid of artery forceps.

To determine whether the cannula had been correctly positioned, the satiated rat was injected with up to 0·01 ml. of 2 per cent. (w/v) NaCl in water. If the cannula tip was lodged within the ventricular system, the injected hypertonic saline had access to the hypothalamic osmoreceptors, and the animal responded by drinking water within 2 min. of the commencement of the injection. The validity of this behavioural test (suggested to me by Prof. J. A. Deutsch) was confirmed several times at autopsy by injecting a dye through the cannula, and any animal that had failed to respond was routinely excluded from further experiment. The rejection rate on this basis was less than one animal in five.

Experiment 1. Satiation of Hunger Drive

It is now well established that lesions of the hypothalamic ventromedial nucleus lead to hyperphagia, while lateral lesions cause aphagia. Electrical stimulation of these areas yields the converse results, but the factors that normally activate these centres to regulate food intake in accordance with nutritional requirements, remain controversial.

Among the several theories now current is the glucostatic hypothesis of Mayer (1955). According to Mayer, the ventromedial nucleus contains receptors sensitive to circulating glucose, which, in the presence of adequate insulin, causes the centre to suppress further intake. Evidence of such receptors has been reported by Marshall, Barrnett and Mayer (1955) who found that intravenous injections of gold-thioglucose in mice led to hyperphagia and to lesions that always involved the ventromedial nucleus. Because of the rapid removal of injected glucose from the blood-stream, direct experimental tests of the glucostatic hypothesis have yielded conflicting results (Bernstein and Grossman, 1956). However, hunger-reducing effects of intraventricular glucose have been described in a preliminary report (Herberg, 1960). It was found that isotonic glucose, with or without added insulin, depressed food intake during test-meals by approximately 35 per cent. ($p < 0{\cdot}001$), while Locke's solution and isotonic fructose were without effect.

Recent work has suggested that glucagon, the hyperglycaemic glycogenolytic secretion of the pancreatic α-cells, may be an important complementary agent. It has been found that intravenous injections of this hormone in man effectively depress gastric hunger contractions and subjective hunger sensations (Schulman, Carleton, Whitney and Whitehorn, 1957), independently of any subsequent rise in blood sugar levels (Stunkard, Van Itallie and Reis, 1955).

Stunkard *et al.* (1955) have suggested that these effects are mediated by the ventromedial nucleus, and that glucagon either acts on the glucoreceptors directly, or increases the receptor utilization of circulating glucose. The following experiment was designed to test these suggestions.

PROCEDURE

Five male hooded rats weighing between 225 and 323 gm. were to be given 10 intraventricular injections each, ordered according to a counterbalanced Latin-square design so that each animal would serve as its own control. To avoid feeding disturbances from gastro-intestinal upsets or injection trauma, no antibiotics were given, and during the second half of the series two animals developed intracranial infections and had to be killed. Only the first five readings were obtained from these two animals.

One of five different sterile solutions was administered at each injection, with an interval of at least 48 hr. between injections. The solutions so tested were: (*a*) A control injection of Locke's solution B.P.C., which contains 100 mgm. glucose/100 ml. (*b*) Locke's solution containing 5·0 μgm. crystalline glucagon/100 ml., which yielded doses of 0·01 μgm. glucagon/kgm. body weight. (*c*) Locke's solution containing 250 mgm. glucose/100 ml. (*d*) Glucose-free Locke's solution containing 500 μgm. glucagon/100 ml., yielding 1·0 μgm. glucagon/kgm. body weight. (*e*) Locke's solution containing 500 μgm. glucagon/100 ml., yielding 1·0 μgm. glucagon/kgm. body weight.

Each injection of 0·02 ml./kgm. lasted between 7 and 13 min. according to the weight of the rat, and commenced at the beginning of a ½ hr. test-meal, when the animal had been starved for 19 hr. At the end of the test-meal the food remaining in the cage was weighed to the nearest 0·25 gm., and the rat was allowed to eat for a further 4½ hr. Water was freely available throughout.

RESULTS

None of the glucagon solutions, with or without 100 mgm. glucose/100 ml., caused any significant fall in intake (see Table I). In the case of the two concentrated glucagon solutions, (*d*) and (*e*), there was a trend in the expected direction, but this may have been due to systemic absorption. Solution (*c*), containing 250 mgm. glucose/100 ml., depressed intake by approximately 8 per cent. ($p < 0{\cdot}02$ using a one-tailed t test).

DISCUSSION

The satiating effect of glucose was again apparent, but it is by no means clearly the sufficient or only factor responsible for activating the ventromedial nucleus. Although the local concentration of glucose following injection (*c*) must have exceeded the upper limits normally found in post-prandial hyperglycaemia, the reduction in intake was only of the order of 8 per cent. Furthermore, the admixture of 0·5 u. soluble insulin per gm. glucose has failed to enhance the anorexigenic action of intraventricular glucose (Herberg, 1960), and these facts emphasize the need for postulating some other synergist, such as glucagon.

The activity of the glucagon supplied had been assayed as causing a rise in blood sugar of 30 to 40 mgm./100 ml. within 10 to 15 min. after an intravenous injection of 0·07 to 0·1 μgm./kgm. The intraventricular doses in the present experiment were considerably smaller than the systemic doses of Stunkard *el al.* and Schulman

TABLE I

GRAMS OF FOOD EATEN DURING ½ HR. TEST-MEALS

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)
Rat	*Locke's solution*	*Glucagon* 5 *μgm./* 100 *ml.*	*Glucose* 250 *mgm./* 100 *ml.*	*Glucose-free glucagon* 500 *μgm./* 100 *ml.*	*Glucagon* 500 *μgm./* 100 *ml.*
1	4·50 (1)	5·25 (5)	4·00 (4)	4·00 (3)	4·50 (2)
	5·75 (10)	5·00 (6)	4·50 (7)	4·25 (8)	5·75 (9)
2	5·75 (2)	5·50 (1)	5·25 (5)	4·50 (4)	5·25 (3)
	—	—	—	—	—
3	7·00 (4)	6·75 (3)	7·25 (2)	7·00 (1)	7·00 (5)
	—	—	—	—	—
4	5·50 (5)	7·00 (4)	5·50 (3)	5·50 (2)	5·75 (1)
	5·00 (6)	4·50 (7)	4·75 (8)	5·75 (9)	4·75 (10)
5	6·00 (3)	5·50 (2)	5·50 (1)	6·00 (5)	6·00 (4)
	5·75 (8)	5·75 (9)	5·00 (10)	6·00 (6)	5·50 (7)
Mean	5·7	5·7	5·2	5·4	5·6

Figures in parentheses indicate the order of injection

et al., which ranged from 14 to approximately 30 μgm./kgm., but solutions (*d*) and (*e*) were calculated to produce local concentrations of glucagon, at the hypothesised site of action, at least equal to the general levels following the systemic administrations.

The failure to observe any significant intake reduction following intraventricular glucagon raises the question whether the glucagon in fact had access to the postulated glucoreceptors. Although the cerebrospinal fluid may be regarded for practical purposes as being largely "co-extensive with the interstitial fluid of the C.N.S." (Tower, 1958), it is possible that the glucoreceptors are located on the blood side of the blood-brain barrier which is relatively permeable to glucose but impermeable to larger molecules and many other compounds, including fructose (Tower, 1958). Such a disposal of the glucoreceptors has already been suggested to account for the ineffectiveness of intraventricular fructose as compared to glucose (Herberg, 1960), but is contrary to the view of Mayer (1953, 1955), who regards the glucosensitive tissue as functionally forming part of the C.N.S. proper.

An alternative interpretation of the present finding is that the anorexigenic action of glucagon is exerted on the hypothalamus indirectly. The intermediary could be some further metabolite or hormone liberated by glucagon from the liver.

EXPERIMENT 2. ACTIVATION OF HUNGER DRIVE

According to the lipostatic hypothesis of Kennedy (1953), the ventromedial nucleus is activated by some intermediary product of fat metabolism. This hypothesis would ensure satiation contingent not only on ingested food but on the organism's current body weight, and Mayer invokes Kennedy's mechanism as a complement to his glucostatic hypothesis so as to account for the long-term regulation of intake. However, the fatty metabolite concerned has not yet been identified.

Long-term regulation might also be achieved by fatty metabolites present and active not during satiety, but during starvation when body fat is being mobilised as fuel. The metabolites would then activate the *lateral* hypothalamus, causing hunger. The non-esterified fatty acids are promising candidates: their blood level has been shown to fall after a mixed meal and to rise during starvation (Dole, 1956; Gordon and Cherkes, 1956).

In this connection, Seifter and Baeder (1956) have isolated from plasma a lipid mobilizing agent which appears in the blood during depletion of the body's carbohydrate stores, or in response to the administration of cortisone. This substance was incidentally found to cause a threefold increase in food intake when injected into rats in single weekly doses, and a series of intraventricular injections were now given to determine whether this effect was mediated by the lateral hypothalamus.

Such duplex control of the ventromedial nucleus and the lateral hypothalamic feeding centres by two mutually independent blood constituents has not been previously considered, but it could theoretically account for the complex feeding pattern induced by ventromedial lesions (Teitelbaum and Campbell, 1958) and for the several refractory observations (Hervey, 1959; Stunkard, 1957) not readily interpretable in terms of any single current theory.

It was found, however, that the mean daily intake of two animals showed no consistent change after intraventricular injections containing 15·0 or 150·0 μgm. Lipid Mobilizer in isotonic saline. The intraventricular doses given were respectively 1/10th and 1/100th of the effective lipaemia-inducing intravenous dose, and are subject to the same considerations discussed in relation to the glucagon administrations.

The above finding does not exclude the possibility that the lateral hypothalamus is sensitive not to Lipid Mobiliser, but to the fatty acids mobilised in response to its secretion. Because of the physical properties of the fatty acids, the intraventricular technique is unsuitable for testing this suggestion.

EXPERIMENT 3. THIRST DRIVE

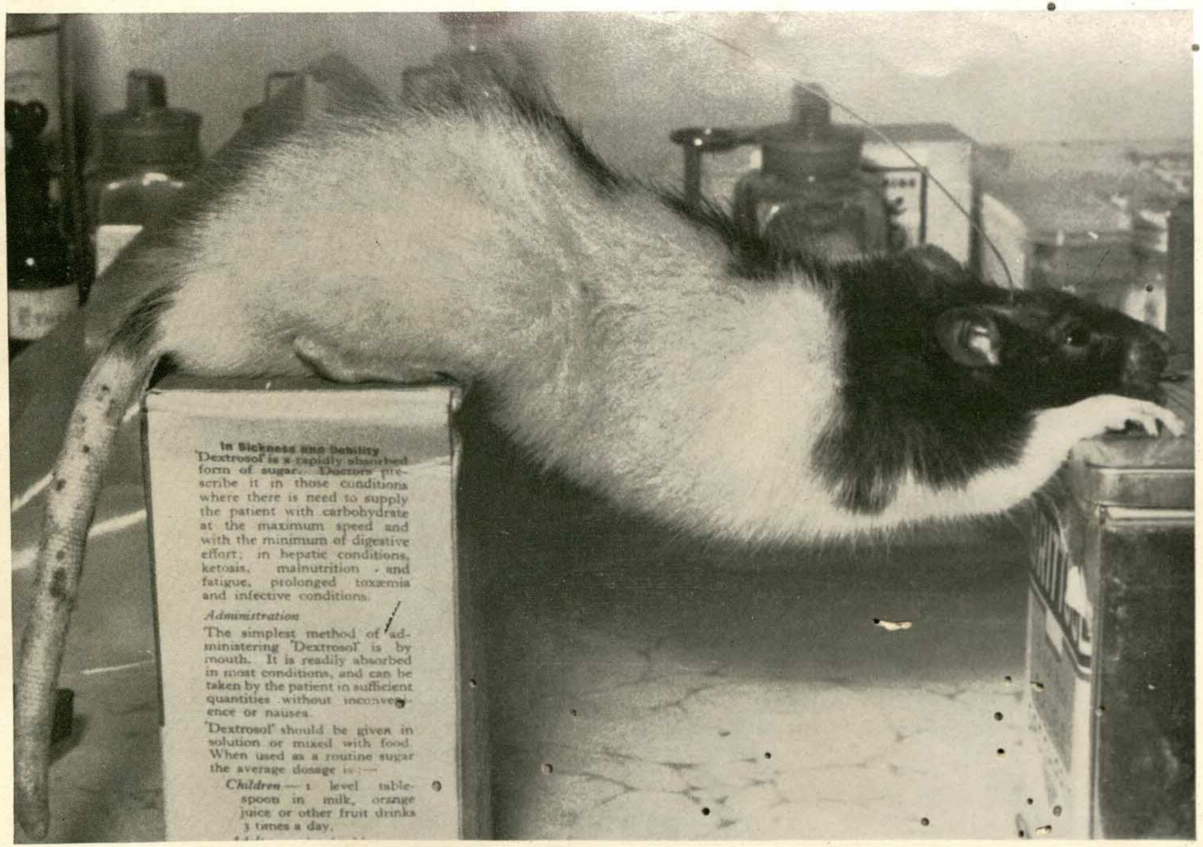
It has already been mentioned that drinking in response to intraventricular 2 per cent. saline in the rat was so dependable that it afforded a simple behavioural test of the correct positioning of the cannula tip. Andersson and Wyrwicka (1957) have shown that the thirst elicited by hypothalamic stimulation in goats has the properties of a genuine primary drive since complex intrumental responses can be conditioned to it.

The converse effect was also confirmed: thirsty rats given intraventricular hypotonic injections (0·02 ml. of 0·75 per cent. NaCl) drank less during a 15 min. test than after isotonic injections. However, this phenomenon did not necessarily involve the normal thirst-quenching mechanisms. Animals usually stop drinking long before any significant dilution of the blood could have taken place, and any internal signals that contribute to the quenching of thirst must thus arise from the oropharynx or stomach. According to theoretical behaviour systems based on stimulus tropisms (Deutsch, 1960), a reduction in the tonicity of the blood or C.S.F. would be thirst-reducing but not rewarding; unlike water by mouth, it could not provide the incentive for performing an instrumental response. This prediction is contrary to neo-behaviourist theory (e.g. Hull, 1952), which equates reward with need-reduction.

EXPERIMENT 4. THE DRIVE FOR SALT

Salt-deprived or adrenalectomized animals develop a salt-craving independently of trial-and-error learning (Epstein and Stellar, 1955). Thus although the nature of the specific appetites is generally obscure, the appetite at least for salt seems to be a primary physiological drive, and it has been suggested that the hypothalamus in some way meters the NaCl of the blood (Andersson and McCann, 1955).

FIGURE 1


In Sickness and Debility
'Dextrosol' is a rapidly absorbed form of sugar. Doctors prescribe it in those conditions where there is need to supply the patient with carbohydrate at the maximum speed and with the minimum of digestive effort; in hepatic conditions, ketosis, malnutrition and fatigue, prolonged toxæmia and infective conditions.
Administration
The simplest method of administering 'Dextrosol' is by mouth. It is readily absorbed in most conditions, and can be taken by the patient in sufficient quantities without inconvenience or nausea.
'Dextrosol' should be given in solution or mixed with food. When used as a routine sugar the average dosage is:—
Children—1 level tablespoon in milk, orange juice or other fruit drinks 3 times a day.

Experimental catatonia following intraventricular procaine. This posture was maintained for nearly 2 min.

To investigate this possibility, six cannulated rats were allowed continuous access to two graduated drinking-tubes, one delivering water and the other a mildly aversive 2–2·5 per cent. (w/v) hypertonic saline solution. Each rat was then given one or more intraventricular injections of 0–0·45 per cent. (w/v) hypotonic NaCl, but no increase in saline intake, immediate or delayed over 24 hr., was found after any of the hypotonic injections. One animal also rejected crystalline salt placed in its cage.

This negative finding is in line with many unsuccessful attempts to relate natriuresis to variations in serum electrolytes *per se* (Smith, 1957). From the point of view of biological economy it is probable that salt-hunger and anti-natriuresis have a common physiological basis, and the present results support the current view that the body's sodium requirements are monitored indirectly, on the basis of the volume of the intravascular fluid compartment.

EXPERIMENT 5. GENERAL DRIVE LEVEL

Feldberg and Sherwood (1954*b*) described a "catatonic state" following the intraventricular administration of acetylcholine, eserine or Di*isop*ropylfluorophosphonate (DFP). The injected cats, though capable of co-ordinated movement (as when pushed off a low table), could be manipulated into a variety of bizarre postures and showed a passive waxy flexibility for long periods. Figure 1 illustrates what appears to be a similar condition in the rat following intraventricular injection of 0·04 ml. 1·0 per cent. procaine hydrochloride. The prodromal signs resembled those vividly described by Feldberg and Sherwood in the cat; vigorous scratching and grooming were again prominent though lacrimation and tremor were not seen.

Procaine is structurally related to the eserine-like anticholinesterases, being a substituted tertiary amine. That the catatonia is not ascribable to a non-specific blocking action by the drugs concerned, is suggested by the absence of similar effects following intraventricular alcohol. In three rats, injections of 4 to 10 per cent. (v/v) ethyl alcohol in Locke's solution were without obvious behavioural effect.

The site of action of the intraventricular anticholinesterases can perhaps be narrowed down to the hypothalamus or the periventricular parts of the limbic system. Experimental lesions of the periaqueductal central grey of the midbrain lead to mutism but not akinesia (Adametz and O'Leary, 1959). Further evidence against a mesencephalic site of action is the inefficacy of intraventricular glucose in hypoglycaemia. After insulin coma had been induced in four rats according to the procedure of Doyle and Olsen (1957), it was found that intraventricuar glucose failed to resuscitate the rats unless given in very high concentrations (e.g. 50 per cent. w/v), and then only after a delay of approximately 20 min., during which interval much of the injected glucose must have entered the blood-stream. Insulin coma is due to the effects of hypoglycaemia on the brainstem reticular formation; the ineffectiveness of intraventricular glucose suggests that the anatomical separation of the reticular formation from the aqueduct and fourth ventricle is too great for prompt diffusion of injected substances from the C.S.F.

The catatonia described by Feldberg and Sherwood is reminiscent of the akinetic mutism sometimes associated with hypothalamic tumours of the third ventricle (Cairns, Oldfield, Pennybacker and Whitteridge, 1941), and it is tempting to regard at least some instances of catatonia as the picture of an organism lacking its hypothalamic drives.

I wish to thank Prof. R. C. Oldfield for laboratory facilities, and the Medical Research Council for providing funds for a stereotaxic instrument. Dr. O. K. Behrens, Lilly Research Laboratories, Indianapolis, kindly supplied the glucagon used in Experiment 1, and Dr. J. Seifter, Wyeth Institute for Medical Research, Philadelphia, the Lipid

Mobilizer used in Experiment 2. This work was done during tenure of the Kenneth Craik Research Award of St. John's College, Cambridge.

REFERENCES

ADAMETZ, J., and O'LEARY, J. L. (1959). Experimental mutism resulting from periaqueductal lesions in cats. *Neurology*, **9**, 636–42.

ANDERSSON, B. (1953). The effect of injections of hypertonic NaCl-solutions into different parts of the hypothalamus of goats. *Acta physiol. scand.*, **28**, 188–201.

ANDERSSON, B., and MCCANN, S. M. (1955). Drinking, antidiuresis and milk ejection from electrical stimulation within the hypothalamus of the goat. *Acta physiol. scand.*, **35**, 191–201.

ANDERSSON, B., and WYRWICKA, W. (1957). The elicitation of a drinking motor conditioned reaction by electrical stimulation of the hypothalamic "Drinking Area" in the goat. *Acta physiol. scand.*, **41**, 194–8.

BERNSTEIN, L. M., and GROSSMAN, M. I. (1956). An experimental test of the glucostatic theory of regulation of food intake. *J. clin. Inv.*, **35**, 627–33.

CAIRNS, H., OLDFIELD, R. C., PENNYBACKER, J. B., and WHITTERIDGE, D. (1941). Akinetic mutism with an epidermoid cyst of the 3rd ventricle. *Brain*, **64**, 273–90.

DEUTSCH, J. A. (1960). *The Structural Basis of Behaviour.* Chicago.

DOLE, V. P. (1956). A relation between non-esterified fatty acids in plasma and the metabolism of glucose. *J. clin. Inv.*, **35**, 150–4.

DOYLE, M. L., and OLSEN, N. S. (1957). Hypoglycaemia in dogs and recovery after intracisternal administration of glucose. *Amer. J. Physiol.*, **191**, 367–70.

EPSTEIN, A. N., and STELLAR, E. (1955). The control of salt preference in the adrenalectomized rat. *J. comp. physiol. Psychol.*, **48**, 167–72.

FELDBERG, W. (1957). Anaesthesia and sleep-like conditions produced by injections into the cerebral ventricle of the cat. *J. Physiol.*, **140**, 20–21*P*.

FELDBERG, W., and SHERWOOD, S. L. (1954*a*). Injections of drugs into the lateral ventricle of the cat. *J. Physiol.*, **123**, 148–67.

FELDBERG, W., and SHERWOOD, S. L. (1954*b*). Behaviour of cats after intraventricular injections of eserine and DFP. *J. Physiol.*, **125**, 488–500.

GORDON, R. S., and CHERKES, A. (1956). Unesterified fatty acid in human blood plasma. *J. clin. Inv.*, **35**, 206–12.

HERBERG, L. J. (1960). Hunger reduction produced by injecting glucose into the lateral ventricle of the rat. *Nature*, **187**, 245–6.

HERVEY, G. R. (1959). The effects of lesions in the hypothalamus in parabiotic rats. *J. Physiol.*, **145**, 336–52.

HULL, C. L. (1952). *A Behaviour System.* New Haven.

KENNEDY, G. C. (1953). The role of depot fat in the hypothalamic control of food intake in the rat. *Proc. Roy. Soc.* B, **140**, 578–92.

MARSHALL, N. B., BARRNETT, R. J., and MAYER, J. (1955). Hypothalamic lesions in goldthioglucose injected mice. *Proc. Soc. exp. Biol. Med.*, **90**, 240–4.

MAYER, J. (1953). Glucostatic mechanism of regulation of food intake. *New Engl. med. J.*, **249**, 13–6.

MAYER, J. (1955). Regulation of energy intake and the body weight: the glucostatic theory and the lipostatic hypothesis. *Ann. N.Y. Acad. Sci.*, **63**, 15–43.

SCHULMAN, J. L., CARLETON, J. L., WHITNEY, G., and WHITEHORN, J. C. (1957). Effect of glucagon on food intake and body weight in man. *J. appl. Physiol.*, **11**, 419–21.

SEIFTER, J., and BAEDER, D. H. (1956). Occurrence in plasma of an extractable lipid mobilizer (LM). *Proc. Soc. exp. Biol. Med.*, **91**, 42–5.

SMITH, H. W. (1957). Salt and water volume receptors. *Amer. J. Med.*, **23**, 623–52.

STUNKARD, A. J. (1957). The effect of the intravenous administration of various nutrients on the gastric hunger contractions of a man with severe brain damage. *Amer. J. clin. Nutr.*, **5**, 203–11.

STUNKARD, A. J., VAN ITALLIE, T. B., and REIS, B. B. (1955). The mechanism of satiety: effect of glucagon on gastric hunger contractions in man. *Proc. Soc. exp. Biol. Med.*, **89**, 258–61.

TEITELBAUM, P., and CAMPBELL, B. A. (1958). Ingestion patterns in hyperphagic and normal rats. *J. comp. physiol. Psychol.*, **51**, 135–41.

TOWER, D. B. (1958). The neurochemical substrate of cerebral function and activity. In Harlow, H. F., and Woolsey, C. N., *Biological and Biochemical Bases of Behaviour.* Madison.

AN INVESTIGATION OF REVERSAL LEARNING IN OCTOPUS VULGARIS LAMARCK

BY

J. MACKINTOSH

From the Institute of Experimental Psychology, University of Oxford

Eighteen octopuses were trained by a method of successive discrimination training to discriminate between a vertical and a horizontal rectangle. They were then overtrained for either 20 or 60 trials, after which they learnt the reverse discrimination. At the beginning of each reversal they were given 20 trials pretraining on the new positive stimulus. The subjects completed from two to nine reversals.

It was found that the amount of overtraining had no effect on the rate of learning the subsequent reversal. The first four reversals all took significantly longer to learn than the original problem, but did not differ significantly from one another. However, examination of the raw data suggests that later reversals do take longer to learn. The performance when learnt (i.e. on the first day of overtraining) was as good after reversal learning as after the learning of the original problem.

The results do not show definite signs that could be interpreted as an exhaustion of a limited supply of neurones available for learning.

INTRODUCTION

Investigations of discrimination reversal learning have been concerned with two aspects of the problem; the effects of repeated reversal, and the effects of overtraining on the original discrimination on the learning of a subsequent reversal. The present experiment on *Octopus vulgaris* Lamarck was concerned with both these aspects.

It has been found that mammals show an improvement in the learning of successive reversals (monkeys—Harlow, 1949; rats—North, 1959; Dufort, Guttman and Kimble, 1954; Pubols, 1957; cats—Cronholm, Warren and Hara, 1960), while animals of other phyla show no such improvement or even a deterioration (pigeons—Reid, 1958; fish—Bitterman, Wodinsky and Candland, 1958; Warren, 1960; an isopod—Thompson, 1957; and a crab—Datta, Milstein and Bitterman, 1960).

In the octopus Young (1956) showed that a single reversal was possible and this was confirmed by Wells and Wells (1957) and Boycott and Young (1958). In addition, Young (1962) investigated repeated reversal; the animals were given 20 trials a day on a black and white discrimination and reversed every day for eight reversals. He found that with successive reversals the number of errors to trials increased. This was due largely to a decrease in the total number of attacks; the attacks that were made were mostly on the positive figure, so that, judged by the number of correct out of total attacks, the performance improved during each day, and perhaps became better on successive days.

To explain his findings Young suggested that learning occurs by the conditioning of new cells. During repeated reversals approximately equal number of cells are conditioned to promote or prevent attack on each of the two figures. This accounts for the fall in the number of attacks. Learning occurs during the course of the day by an increase in the number of cells conditioned to promote attack on the positive figure.

He also found that performance was more accurate when the figure originally positive (black) was again positive, and suggested that this could be explained by a previous theory (Young, 1961), that the supply of cells available for conditioning may become exhausted; all the cells responsible for encoding "black" may have been conditioned to produce attack during the initial learning.

An alternative explanation of the tendency to stop attacking is that as the octopuses in Young's experiment were never trained to a learning criterion, they made a large number of errors in the early reversals and received too many shocks. Therefore, in the present experiment, although punishment was used, precautions were taken to minimize the number of shocks given. A simple visual discrimination was used—a horizontal versus a vertical rectangle (Sutherland, 1957, has shown this to be easy for the octopus); training was continued until a learning criterion was reached; the animals were pretrained to attack the new positive shape for 2 days before each reversal, and a 5-volt shock (in place of the 11-volt shock used by Young) was given.

Training to a criterion before reversal was of further interest because, if the learning mechanism operates in the way Young suggests, it would be expected that the more thorough the learning the more rapidly would the supply of cells available for learning be used up.

An additional purpose of the present investigation was to see whether overtraining has an effect on octopus. It has been found that the rat shows improved learning of the reversal after overtraining (Reid, 1953; Pubols, 1956; Capaldi and Stevenson, 1957), while birds (Warren, Brookshire, Ball and Reynolds, 1960) and fish (Warren, 1960) are hindered by overtraining. Consequently, before each reversal, half the animals were overtrained by 20 trials and half by 60.

PROCEDURE

Experimental design

The subjects were divided into two major groups receiving different amounts of overtraining on each reversal problem. Within each group one half of the subjects were first trained on the vertical rectangle (the preferred shape) as positive, the other on the horizontal rectangle as positive, so that there were four groups as follows:—

Group A: 60 trials overtraining, vertical positive, horizontal negative in original training.
Group B: 60 trials overtraining, horizontal positive, vertical negative in original training.
Group C: 20 trials overtraining, vertical positive, horizontal negative in original training.
Group D: 20 trials overtraining, horizontal positive, vertical negative in original training.

There were 20 trials a day and the criterion used was 80 per cent. correct responses on any one day's trials.

Subjects

The subjects were 18 *Octopus vulgaris* Lamarck obtained from the Bay of Naples. They weighed from 150 to 500 gm. They were kept in individual tanks filled with continuously circulating sea-water. At one end of the tank was a brick "home" in which the animal sat.

Acclimatization

The subjects were left alone in their tanks for two or three days after being caught. Then pieces of sardine on a prong were offered 10 times a day at the end of the tank away from the "home" until they were taken regularly within 5 sec.

Pretraining

The positive shape was then presented 10 times a day. If subjects failed to attack it within 20 sec., a piece of sardine was shown and they were allowed to attack both together. Pretraining was continued until the shape was attacked within 20 sec. at each presentation.

Training

The subjects were then trained on a successive discrimination between vertical and horizontal rectangles of 10 × 2 cm. side length. There were 20 trials a day in two blocks of ten. (For further details of training octopus see Sutherland, 1957.)

After the training criterion of 80 per cent. correct responses on any one day's trials had been reached, reversal was begun. The subjects were first pretrained on the new positive shape (i.e. the previous negative) for 2 days. There were 10 trials a day at each of which the shape was presented alone for 20 sec. and, if not attacked within this time, with a piece of sardine until the subjects attacked both together. It should be noted that no criterion of 10 out of 10 attacks on the shape alone on any one day's trials—as was used in the original pretraining—was used in the reversal pretraining. After the 20 pretraining trials reversal training was started and continued until the previous training criterion of 80 per cent. correct responses was reached.

RESULTS

The data presented are calculated from the number of attacks required to learn a discrimination and not the number of trials. The number of trials shows the

TABLE I

THE NUMBER OF ATTACKS NEEDED TO REACH CRITERION BY INDIVIDUAL SUBJECTS FOR SUCCESSIVE REVERSAL PROBLEMS

Group	*Ro*	*R1*	*R2*	*R3*	*R4*	*R5*	*R6*	*R7*	*R8*	*R9*	*R10*
A (Vertical positive in Ro. 60 trials overtraining)	25	26	43	46	36	—	—	—	—	—	—
	24	48	22	26	48′	—	—	—	—	—	—
	27	41	19	95	103′	—	—	—	—	—	—
	14	64	33	22′	—	—	—	—	—	—	—
	14	37	41	—	—	—	—	—	—	—	—
B (Horizontal positive in Ro. 60 trials overtraining)	20	15	20	29	27′	—	—	—	—	—	—
	30	33	30	89	—	—	—	—	—	—	—
	46	43	21	60	42	—	—	—	—	—	—
	21	26	23	56	62	—	—	—	—	—	—
C (Vertical positive in Ro. 20 trials overtraining)	10	40	25	52	25	55	38	50	48	—	—
	13	17	15	23	30	49	46	48	59	51	43′
	13	53	26	31	24	72	55	132′	—	—	—
	10	68	74	64	72	—	—	—	—	—	—
	13	38	30	60	45	56′	—	—	—	—	—
D (Horizontal positive in Ro. 20 trials overtraining)	47	19	33	—	—	—	—	—	—	—	—
	21	51	61	36	34	—	—	—	—	—	—
	29	45	59	—	—	—	—	—	—	—	—
	30	41	74	47	43′	—	—	—	—	—	—
Averages	*Ro*	*R1*	*R2*	*R3*	*R4*	*R5*	*R6*	*R7*	*R8*	*R9*	*R10*
A	21	43	32	47′	62′	—	—	—	—	—	—
B	29	29	23	58	43	—	—	—	—	—	—
C	14	43	34	46	39	58′	46	77	51	51	43′
D	32	39	57	41	38′	—	—	—	—	—	—
A and B	25	37	28	53′	53′	—	—	—	—	—	—
C and D	22	41	44	45	39′	—	—	—	—	—	—
A, B, C and D	24	39	38	49′	49′	—	—	—	—	—	—

(Note; Ro is original training; ′ indicates experiment stopped before criterion reached.)

same trend as the number of attacks but is larger absolutely and in some cases relatively as well. The number of attacks is a fairer measure of the speed of learning in a successive discrimination situation, since an animal which stops attacking completely will not have an opportunity to learn (*v.* Sutherland, 1960 (p. 11), for further discussion of this point).

Table I shows the number of attacks taken by individual subjects to reach the criterion on successive reversal problems. (Ro is the original training). Subjects completed from two to nine reversals.

The effect of overtraining

Groups which had received different amounts of overtraining were compared by *t*-tests. Tests were not performed where differences in rates of learning were obviously not significant. Only in one case was a significant difference found; in R2, Group B, overtrained by 60 trials, reverse more quickly than Group D, overtrained by 20 trials ($t = 3$, $d.f. = 3$, $p < 0{\cdot}05$). However, since there is no significant difference between the A and C groups for the same reversal, and since no group was consistently faster or slower than the others throughout, it appears that the one significant difference obtained was due to factors other than overtraining, and that overtraining does not affect the speed of reversal.

Successive reversals

Tables II and III compare various quantities for successive reversal problems; *t*-tests were only performed where some data for subjects from all four groups were available (data for unfinished animals were included in the comparisons of R4 since three of these unfinished animals took more attacks to learn R4 than R2—two taking considerably more).

TABLE II

Performance on the First Day of Overtraining

Reversal	*Number of subjects*	*Per cent. correct responses*
Ro	18	85·0
R1	18	83·9
R2	15	81·0
R3	11	87·7
R4	5	84·0
R5	3	80·0
R6	3	86·7
R7	2	92·5
R8	1	90·0
R9	1	80·0

(*t*-tests comparing Ro with R1 to R4 gave values of *t* which were not significant.)

When comparing the learning of successive reversals it is important to compare *alternate* rather than *adjacent* reversals with each other, since the preferred shape was successively the positive then the negative figure. It was found that, although animals took longer to learn R1 than Ro ($t = 3{\cdot}4$, d.f. $= 17$, $p < 0{\cdot}01$), they did not take significantly longer to learn subsequent reversals up to R4 than they did to learn R1. Also R4 did not require more attacks to learn than R2 (its proper comparison problem). Examination of the raw data, however, does suggest there is a tendency for learning to take longer; 10 subjects take longer to learn R3 than R1

TABLE III

The Performance on the last day of pretraining and the First day of Reversal Training

(*a*) Group Averages

Reversal	*Number of subjects*	*Number of attacks on last day of pretraining (out of 10)*	*Number of attacks on first day of reversal training*		
			Total	*Positive*	*Negative*
R0	18	10·00	13·7	9·2	4·5
R1	18	5·64	11·6	5·6	6·0
R2	18	8·28	13·1	5·9	7·2
R3	15	7·00	11·8	5·9	5·9
R4	12	6·83	10·8	5·2	5·6
R5	4	9·50	17·5	9·0	8·5
R6	3	9·00	15·0	8·3	6·7
R7	3	7·00	16·0	8·0	8·0
R8	2	9·00	16·5	8·0	8·5
R9	1	10·00	15·0	6·0	9·0
R10	1	9·00	16·0	8·0	8·0

(*b*) *t*-Tests of Significance

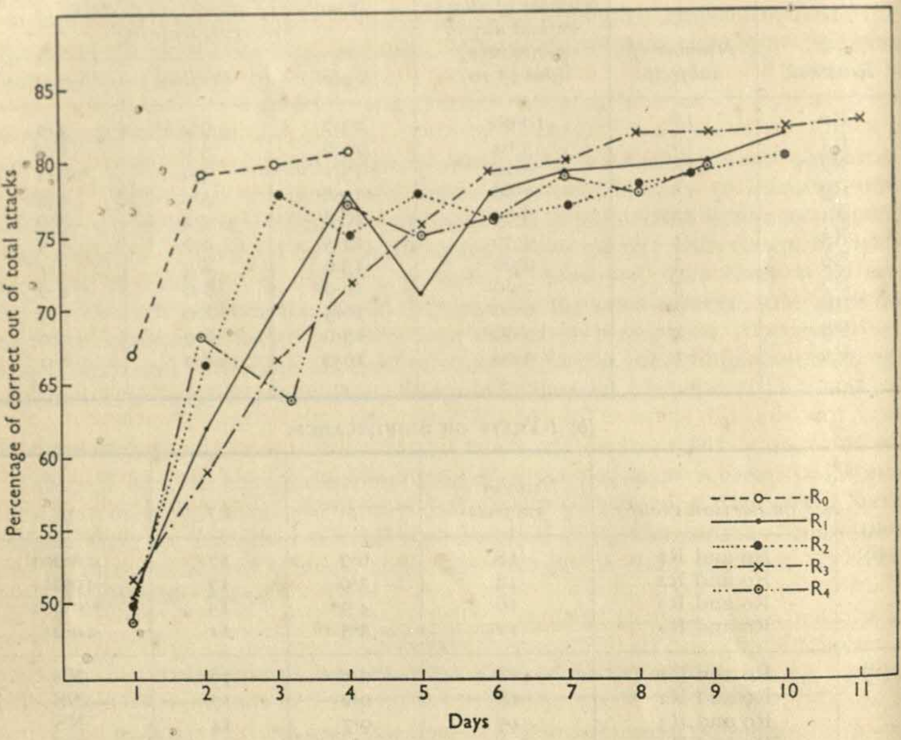
	Comparison groups	*Number of subjects*	*t*	*d.f.*	p
(i)	R0 and R1	18	6·7	17	<0·001
	R0 and R2	18	3·9	17	<0·01
	R0 and R3	16	4·9	15	<0·001
	R0 and R4	12	3·5	11	<0·01
(ii)	R0 and R1	18	2·1	17	NS
	R0 and R2	18	0·2	17	NS
	R0 and R3	15	0·7	14	NS
	R0 and R4	12	0·7	11	NS
(iii)	R0 and R1	18	5·1	17	<0·001
	R0 and R2	18	3·4	17	<0·01
	R0 and R3	15	2·0	14	NS
	R0 and R4	12	3·8	11	<0·01
(iv)	R0 and R1	18	1·7	17	NS
	R0 and R2	18	1·6	17	NS
	R0 and R3	15	1·8	14	NS
	R0 and R4	12	2·0	11	NS

(i) Attacks on positive on last day of pretraining.
(ii) Attacks on positive and negative on first day of reversal.
(iii) Attacks on positive on first day of reversal.
(iv) Attacks on negative on first day of reversal.

and only four are quicker, and similarly eight take longer to learn R4 than R2 and only three are quicker. Also, if the earlier and later reversals for the two animals in Group C who completed eight or more reversals are compared, it can be seen that later reversals take longer (for animal C2, R2 and R4 take a total of 45 attacks while R6 and R8 take 105; R3 and R5 take 72 attacks and R7 and R9 take 99; for animal C1, R2 and R4 take 50 attacks to learn and R6 and R8 take 86).

Once the problem has been learned the performance (i.e. on the first day of over-learning) is just as good following the reversal learning as the original learning (see Table II).

FIGURE 1



Graphs showing the percentage of correct out of total attacks on each day until the criterion was reached for reversals Ro to R4 for all animals combined.
(Data for the graphs were obtained by separately adding the numbers of correct and total attacks for every animal on successive days of training. Where some animals completed their training before others they were included on the assumption that they would continue to score at the 80 per cent. correct level.)

An examination of the learning curves shown in Figure 1 shows that one of the reasons why reversals R1 to R4 take longer to learn than Ro is because the curves begin at a lower point. Table III makes clear that this is because the positive shape is attacked less on the first day of the reversals than on the first day of the original training. An increase in the number of attacks on the negative partly compensates for this but the increases are not significantly greater than the number on the first day of original training. The decrease in attacks on the positive on the first day of learning is clearly correlated with the decrease in the number of attacks on the last day of reversal pretraining. In view of this it seems probable that had the subjects been reversal pretrained to a criterion of 10 out of 10 attacks on the positive during one day's pretraining the learning curves would all have begun at the same point. It is necessary to stress the importance of the number of attacks because, in a successive discrimination, failure to react to the stimuli leads to loss of the opportunity to learn the significance of the stimuli.

Further examination of the learning curves reveals that there is little difference between Ro and the others in the slope of the first part, from day 1 to day 2. However, when a comparison is made between parts of the curves from one level of learning to another—e.g. from 70 to 80 per cent. correct—the Ro curve is steeper than any of the others.

Attention has already been drawn to the fact that alternate reversals are more properly comparable with one another than adjacent reversals due to the facilitation of learning when the preferred shape is positive. When individual animals results are consulted, a "see-saw" effect in learning the odd and even-numbered reversals is more apparent in some cases than in others. Group C animals show the greatest "see-saw" effect; animals C1, C2, C3 and C5 all learn the even-numbered reversals more quickly than the odd-numbered reversals. Some effect can also be seen in Group A, e.g. in animals A2 and A3. In the other two groups, B and D, the effect is far less consistent. While animal D4 learns reversals with the preferred shape positive more quickly than the other reversals, animal B4 learns the *non*-preferred shape positive reversals more quickly.

It was not possible to keep a complete record of behaviour changes throughout the experiment but some observations were made on changes in the attacking behaviour. These changes all suggested caution: animals sometimes darted back after touching the shape or touched it several seconds after approaching it; often the transparent rod to which the shape was attached, rather than the shape itself, was attacked. Cautious behaviour is shown also in difficult discriminations. In ten of the subjects a peculiar "conflict" posture was seen. The animals left the home and swam out until they were half-way between the home and the shape. Then the body became very pale and elongated and the skin developed warty papillae. The arms took on a crumpled appearance owing to the twisting in opposite directions of adjacent sucker-bearing portions of their ventral surfaces. The whole body had a curious rigidity and there was a gentle swaying movement to and away from the shape. Sometimes this posture was abandoned and the shape taken and sometimes it was maintained for several seconds after the removal of the shape. This "conflict posture" has not been recorded before though this does not mean it does not appear in other situations.

Discussion

The effects of overtraining

The negative finding that overtraining does not significantly affect reversal learning suggests, though does not establish, that the octopus is unique among the animals so far investigated. Although the number of overtraining trials given was not absolutely very large (20 or 60) it was relatively large compared with the number of trials required to learn (20 or 40); if overtraining has any effects they should have been apparent.

Successive reversals

The failure to find improvement in the learning of successive reversals is in keeping with the general picture emerging from experiments with diverse taxonomic groups; only some mammals learn later reversals more rapidly than the original problem while the other animals investigated show no such improvement.

Unfortunately the present results do not show conclusively whether or not reversals after the first deteriorate or remain at the same level, and the problem is complicated by the level of attack on the first day of reversal training not being the same. However, the balance of evidence suggests that successive reversals do become progressively, though not dramatically, more difficult to learn.

The results do not show definite signs that could be interpreted as an exhaustion of a limited supply of neurones. However, the finding that performance during overtraining, once the discrimination has been learnt, does not deteriorate with progressive reversal is against rather than for a neurone-store theory, postulating that neurones are distributed to opposing response "blocks." Such a theory should surely predict a growing instability of performance as learning proceeds by the recruitment of an ever-diminishing store of neurones into two ever-increasing rival response factions.

This work forms part of a project on "Stimulus Analysing Mechanisms": the project is financed jointly by the American Office of Naval Research (Contract N62558–2453) and the Nuffield Foundation. The experimental work was carried out at the Stazione Zoologica, Naples during the tenure of the Naples Biological Scholarship awarded by Oxford University. I would like to thank the director and staff of the Stazione for the hospitality and facilities afforded me there, and Dr. N. S. Sutherland and Professor J. Z. Young for much advice and encouragement.

REFERENCES

Bitterman, M. E., Wodinsky, J., and Candland, D. K. (1958). Some comparative psychology. *Amer. J. Psychol.*, **71**, 94–110.

Boycott, B. B., and Young, J. Z. (1958). Reversal of learned responses in *Octopus vulgaris* Lamarck. *J. anim. Behav.*, **6**, 45–52.

Capaldi, E. J., and Stevenson, H. W. (1957). Response reversal and amount of overtraining. *J. comp. physiol. Psychol.*, **50**, 195–8.

Cronholm, J. N., Warren, J. M., and Hara, K. (1960). Distribution of training and reversal learning by cats. *J. genet. Psychol.*, **96**, 105–13.

Datta, L. E. G., Milstein, S., and Bitterman, M. E. (1960). Habit reversal in the crab. *J. comp. physiol. Psychol.*, **53**, 275–8.

Dufort, R. H., Guttman, N., and Kimble, G. A. (1954). One-trial discrimination reversal in the white rat. *J. comp. physiol. Psychol.*, **47**, 248–9.

Harlow, H. F. (1949). The formation of learning sets. *Psychol. Rev.*, **56**, 51–62.

North, A. J. (1959). Discrimination reversal with spaced trials and distinctive cues. *J. comp. physiol. Psychol.*, **52**, 426–9.

Pubols, B. H. (1956). The facilitation of visual and spatial discrimination by overlearning. *J. comp. physiol. Psychol.*, **49**, 243–8.

Pubols, B. H. (1957). Successive discrimination reversal learning in the white rat: a comparison of two procedures. *J. comp. physiol. Psychol.*, **50**, 319–22.

Reid, L. S. (1953). The development of non-continuity behaviour through continuity learning. *J. exp. Psychol.*, **46**, 107–12.

Reid, R. L. (1958). Discrimination-reversal learning in pigeons. *J. comp. physiol. Psychol.*, **51**, 716–20.

Sutherland, N. S. (1957). Visual discrimination of orientation by octopus. *Brit. J. Psychol.*, **48**, 55–71.

Sutherland, N. S. (1960). Visual discrimination of orientation by *Octopus*: mirror images. *Brit. J. Psychol.*, **51**, 9–18.

Thompson, R. (1957). Successive reversal of a position habit in an invertebrate. *Science*, **126**, 163–4.

Warren, J. M. (1960). Reversal learning by paradise fish (*Macropodus opercularis*). *J. comp. physiol. Psychol.*, **53**, 376–8.

Warren, J. M., Brookshire, K. H., Ball, G. G., and Reynolds, D. V. (1960). Reversal learning by white leghorn chicks. *J. comp. physiol. Psychol.*, **53**, 371–5.

Wells, M. J., and Wells, J. (1957). The effect of lesions to the vertical and optic lobes on tactile discrimination in *Octopus*. *J. exp. Biol.*, **34**, 378–93.

Young, J. Z. (1956). Visual responses by octopus to crabs and other figures before and after training. *J. exp. Biol.*, **33**, 709–29.

Young, J. Z. (1961). Learning and discrimination in the octopus. *Biol. Rev.*, **36**, 32–96.

Young, J. Z. (1962). Repeated reversal of training in *Octopus*. *Quart. J. exp. Psychol.* (in press).

BINOCULAR RIVALRY AND STEREOSCOPIC DEPTH PERCEPTION

BY

ANNE TREISMAN

From the Institute of Experimental Psychology, University of Oxford

An investigation was made of stimulus factors causing retinal rivalry or allowing stereoscopic depth perception, given a requisite positional disparity. It is shown that similar colour information can be "filtered" out from both eyes; that stereopsis is not incompatible with rivalry and suppression of one aspect of the stimulus, and that the strongest cue for perception of stereoscopic depth is intensity difference at the boundaries of the figures in the same direction at each eye. Identity of colour can also act as a cue for stereopsis. The brightness of different monocular figures seen in the stereoscope in different combinations was estimated by a matching technique, and it is suggested that the perceived brightness is a compromise between the monocular brightness difference between figure and ground seen in relation to the binocular fused background, and the mean brightness of the figures. The results are discussed in terms of neurophysiological "on," "off" and continuous response fibres.

INTRODUCTION

There has been relatively little investigation of binocular rivalry and suppression. Little is known of what aspects of the stimulus information are inhibited under different conditions, whether different areas are suppressed at random in rivalry, or the input from one eye as a whole, or whether any other form of "filtering" is possible. This contrasts with the considerable body of research on the aspects of auditory messages which can selectively evoke responses. Two recent papers have described a selective effect of interest or motivation on alternations in rivalry: Engel (1956) using inverted and upright photographs of faces and Bagby (1957) with typical culture-related scenes; these suggest selective control from other parts of the brain, and involved suppression of one eye as a whole. Bárány and Hallden (1948) have shown that suppression of one eye may take place relatively peripherally, finding that the pupillary reflex of the eye which is temporarily not perceiving may be suppressed during rivalry.

To describe this group of findings the term "retinal" rivalry is appropriate, but it would be interesting to consider whether this is the only non-random form of suppression. Certainly, if two different stimuli are presented to non-corresponding areas of the two retinae, they may be seen side by side; if the two whole fields are different, the resulting perception may be of a shifting, changing patchwork. What criteria determine whether the inputs from corresponding points of the two eyes are sufficiently similar to fuse, or sufficiently different to compete? It is generally agreed, since Breese's classical findings (1909), that fields of widely different colours may rival, and that different contours on corresponding areas almost certainly will do so.

This paper will be concerned mainly with the following questions: can the brain select different signals of one colour presented to both eyes, while at the same time rejecting signals of another colour also presented to both eyes? In other words is rivalry always between retinae, or can it also occur between categories of stimulus information? Secondly, is one aspect of a single stimulus, such as its colour, sometimes suppressed, while another aspect, such as the position of the figure, is retained?

One of the few reports to throw any light on this point is an interesting paper of Creed's (1935). He presented postage stamps of different colours and slightly different designs to each eye, and found that the design of one stamp was sometimes seen in the colour of the other. The method used in the present experiment was to see whether stereoscopic depth perception was retained when stimuli (with the requisite position disparity) were made sufficiently incompatible in another dimension to produce rivalry and suppression. For instance, if they were in complementary colours, so that only one colour was seen at a time, did this mean that all the information about the suppressed coloured figure was lost, or could its shape and position be retained to give rise to a stereoscopic impression? Similarly, different brightnesses of the figures provided potentially incompatible inputs, which again might be reconcilable with stereopsis.

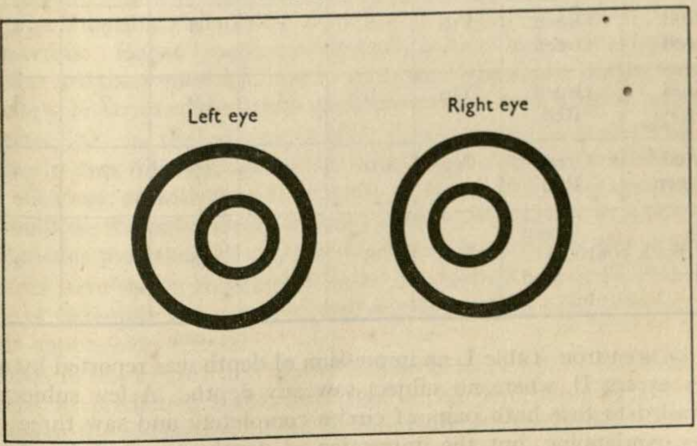
Arising from these experiments came the more specific question: what features of the stimulation of slightly disparate retinal points will give rise to the perception of a fused, three-dimensional figure? A considerable body of evidence has accumulated on the minimum and maximum fusional areas, on the horopter, and generally on the position information providing the basis for stereopsis (Ogle, 1950; Rønne, 1956); but little is known about other aspects of the stimuli which provide the necessary conditions. Rønne, for example, in his recent review, says, "the coupling of disparate cells only takes place if the inputs are identical, in which case the percept is visualized in front (or behind) the horizon." It is the limits or modifications to this "identity" of the inputs which will be discussed in the present paper.

The fact that it appears to be contours which cause most rivalry or suppression of competing areas on the two retinae may also give a clue to the type of compatibility of signals necessary for stereopsis. Contours or boundaries of figures consist of step-changes in intensity in the visual field, and two possible neurophysiological systems that may be partly responsible for signalling them are: (1) the firing of "on" and of "off" fibres, generated by eye movements making the discontinuities in illumination cross different retinal cells. Ditchburn (1955) has shown that when eye movements are eliminated, contours rapidly fade and disappear. (2) Hubel and Wiesel (1959), recording from cells in the cat's striate cortex, found receptive fields composed of central areas giving either "on" or "off" responses and flanking areas with the opposite excitatory or inhibitory effects; they say "the exact position of the vertical border between light and darkness was very critical for a strong response."

One possibility this suggests is that rivalry for shapes or patterns may be due to conflicting "on" and "off" signals from corresponding points of the two eyes. If this were so, one might predict that figure-ground boundaries consisting of step changes in intensity in the same direction from the background to the figure on each eye would be acceptable for "fusion" and give rise to an impression of stereoscopic depth, while those in opposite directions would not, even though they lay within the stereoscopic fusional areas. This again would be consistent with Hubel and Wiesel's reports of cortical cells with receptive fields of the same shape in homologous areas of both eyes. These fields had the same excitatory and inhibitory characteristics in each retina, and showed the same principles of summation and antagonism which had been demonstrated for fields in the single eye. In the behavioural field, Fry and Bartley (1933) had earlier published some related findings, and came to the conclusion, as Bartley (1941) summarizes it, that "binocular rivalry between the processes in the two eyes set up by stimulation from A, occurs only when the one is above and the other below the brightness of the ground." One could take these findings as some support for the view that "fusion" depends on correlated firing from the "on" and "off" fibres in the two eyes. It would be interesting to determine if this were also a condition of stereoscopic depth perception, and, if so, whether it is the only cue used, or whether figures of the same colour but with an opposite intensity difference from their background, or no intensity difference at all, will also give rise to an impression of depth.

The use of systematically varied stimuli in the stereoscope could also throw more light on whether certain visual effects occur at the peripheral and monocular level or depend on the more central binocular systems. For example, does simultaneous contrast of brightness and colour occur in relation to the monocular background or to the perceived, "fused," binocular background? If it is the former, can induced contrast interfere with stereopsis and cause rivalry? An experiment by Asher (1950) who found no difference in brightness contrast with stimuli projecting to the same and to opposite sides of the cerebral cortex from one eye, provides some evidence that phenomena of contrast do originate in the retina. Asher also found that monocular brightness judgements showed considerable independence; a dim light in one eye on a dark ground appeared bright by contrast, and retained this heightened brightness even when the background against which it was perceived was a much more intensely illuminated field received by the other eye. There is room for more investigation of how these dark-light relationships combine information from the two eyes. The well known Fechner paradox stated that a weak light in one eye, instead of summating with a brighter one in the other, makes it look dimmer, the perceived brightness being some compromise between the two; does this also hold with comparative judgements of brightness, for instance of figure and ground? Fry and Bartley (1933) have shown that this averaging of brightness varies not only with the absolute brightness of the dimmer stimulus, but also with the difference between it and the ground, and have suggested that the effects of binocular summation and inhibition combine. Similar problems arose independently in the course of the investigation to be reported here, and related findings will be discussed.

FIGURE 1



EXPERIMENTAL METHOD AND RESULTS

A pair of annuli (Fig. 1) were used as the stimuli in all the conditions to be described. They were 2·5 mm. wide and filled in either in black, white, grey or some colour. Their diameters were 35 and 16 mm., and one was inside the other with its centre displaced by 2·5 mm. in the nasal direction. The circles were cut out of coloured paper and stuck on to cards 7 cm. high and 4 cm. wide. They were presented in an Asher–Law stereoscope at an adjustable distance from the eyes, varying from 15 cm. to 25 cm., so that the disparity varied from about 0·5 to 0·9 degrees. The cards were illuminated from above by a 100-watt anglepoise lamp. Subjects, after adjusting the focal distance, were asked to describe what they saw and were questioned on any points of interest which they omitted.

(1) *Colour rivalry and stereopsis*

The first series of cards used consisted of red and green circles on a white ground. Subjects were familiarized with the perception of stereoscopic depth with card A which had red circles on both eyes. These were always seen as clearly in different planes. They were then tested with the other cards described in Table I, all of which had signals in different colours to the corresponding areas of each eye. The two main points of interest were whether the cards would give an impression of depth despite the colour opposition, and whether rivalry between the colours would take the form of suppression at random (patchy appearance), suppression of one eye, or suppression of one colour. With cards C and D, the latter would have involved selecting one circle from each eye. Card E was included to see how a compatible contour (black in both eyes) round the inner and outer edges of the annuli affected the stereopsis and the red-green rivalry. Card D, which had no positional disparity, was a control, to ensure that subjects were not reporting depth when there was no stereoscopic cue provided.

TABLE I

						Colours			
Card	*Left eye circles*	*Right eye circles*	*Nasal dis-parity*	*Depth*	*Dip-lopia*	*Patchy*	*Both one colour*	*Circles oppte. colours*	*No. of S's*
A	Red Red	Red Red	Yes	8	0	0	8	0	8
B	Red Red	Green Green	Yes	8	1	8	7	1	8
C	Red Green	Green Red	Yes	9	2	9	7	4	9
D	Red Green	Green Red	No	0	0	7	7	4	8
E	Red (Black contour) Green	Green Red	Yes	8	3	8	2	2	8

As can be seen from Table I, an impression of depth was reported by all subjects in all cases except D, where no subject saw any depth. A few subjects on some occasions failed to fuse both pairs of circles completely and saw three circles, the centre ones overlapping, but the impression of depth remained despite the slight diplopia. (Burian, 1936, has reported a similar retention of depth despite some diplopia.)

The colours in this series were never seen as mixed (yellow, grey, or muddy coloured). They were seen either as patchy, fluctuating and "chasing each other round," or as alternately one colour and then the other, or as the two whole circles in different colours. Many subjects often and some occasionally saw both circles in the same colour even with cards C and D, where this meant seeing one circle from one eye and one from the other. This was, if anything, more frequent than seeing the circles one completely green and one completely red, although this would have been true "retinal" rivalry. Depth was perceived despite this suppression of half the colour information. The single colour was much more rare with the black contours round the annuli (card E), which were described as much more unstable and patchy, although the depth effect appeared to be neither more nor less pronounced.

It seemed then that information about the colour could be suppressed in rivalry while the position and shape of the stimulus were retained. Since the disparity cue for depth only exists if information from both eyes is used, and since the circles were sometimes seen all in one colour with card C, it is safe to say that competing signals were not suppressed here (as in Bárány's experiment) on one retina as a whole, and that even when a signal of one colour was apparently suppressed, information about its position or contour was being used.

(2) *Intensity differences, rivalry and stereopsis*

A second, more detailed series of experiments was carried out to find more exactly what aspects of the signals were being compared and functioning as similar enough to give rise to stereopsis. The method used was to vary the direction of change in intensity at the boundaries between figure and ground. A number of cards were prepared, varying the brightness of both circles and backgrounds and also their colours. The intensities of the light reflected by circles and background were measured in log foot lamberts, using an S.E.I. photometer, in the experimental situation. (Each measurement was the mean of ten readings, and its standard error was approximately 0·01 log foot lamberts.) These intensities are given in Table II, together with the description of each card and the results on presence of depth and of diplopia. Since the hypothesis was that the direction of the intensity difference between figure and ground would be relevant, this is given as well as its extent. The absolute values of the intensity differences given in the table are roughly comparable in subjective terms. Hanes (1949), for example, gives a scale of subjective brightness (very similar to the cumulated j.n.d. scale for brightness) which appears to be approximately linear over the range used here, of 0·4 to 1·6 log foot lamberts.

The letter "O" in the column headed "circle colours" means that where the circles were in two different colours or brightnesses, there was one circle of each colour on each eye, as with card C in the first series. Thus the upper entry in that column would be the outer circle on one eye and the inner one on the other. The letter "S" means that the pair on one eye were the same colour or brightness.

The cards were shown to subjects in the stereoscope as before, and the presence or absence of stereopsis was noted, as well as other details such as fusion or diplopia, and colour suppression effects.

The cards in black, grey and white without colours will be discussed first. The results were, as predicted, stable and clear depth perception by all subjects with card X, where the circles were both lighter than the background, while the same circles on an intermediate grey background, card Y, were seen with stereopsis only by one subject occasionally. In card Z, the difference in intensity in opposite directions on each eye was achieved by placing identical grey circles, one pair on a white card for one eye and the other on black for the other. Here again only one subject obtained the impression of stereoscopic depth, while all subjects saw two rivalling pairs of light and dark grey circles. The difference was due, presumably,

TABLE II

Card	Circle colours	Log foot lamberts	Back-ground colours	Log foot lamberts	Intensity differences	Depth perceived		Diplopia		No. of S's
						Always	*Occa-sionally*	*Always*	*Occa-sionally*	
X	Dark grey White (O)	0·80 1·53	Black	0·43	+0·37 +1·10	13	0	0	0	13
Y	Dark grey White (O)	0·80 1·53	Light grey	1·21	−0·41 +0·32	0	1	10	1	14
Z	Dark grey Dark grey	0·80 0·80	White Black	1·53 0·43	−0·73 +0·37	1	0	9	1	14
XY	Dark grey White (O)	0·80 1·53	Black Light grey	0·43 1·21	{+0·37, {−0·41 +1·10, +0·32 Left eye Right eye	5	6	5	3	13
P	Red Green (O)	1·08 1·02	White	1·53	−0·45 −0·51	11	0	2	2	11
Q	Red Green (O)	1·08 1·02	Black	0·43	+0·65 +0·59	7	3	0	2	10
R	Red Green (O)	1·08 1·02	Grey	1·06	+0·02 −0·04	2	4	7	0	12
S	Yellow Blue (O)	1·45 0·81	White	1·53	−0·08 −0·72	1	6	3	1	10
T	Yellow Blue (O)	1·45 0·81	Black	0·43	+1·02 +0·38	7	3	3	0	11
U	Yellow Blue (O)	1·45 0·81	Grey	1·06	+0·39 −0·25	0	0	7	4	12
G	White White	1·53 1·53	Green Red	1·02 1·08	−0·51 −0·45	10	1	0	0	11
H	Red Green (S)	1·08 1·02	Green Red	1·02 1·08	+0·06 −0·06	2	6	4	2	14

Significance of χ^2 comparisons between cards in Table II, on presence or absence of stereoscopic depth perception.

In order to obtain sufficiently large expected values in each cell to make χ^2 meaningful, some of the cards had to be grouped and in most of the calculations the cell for "occasionally" judgements was grouped with the "always" or with the "never" judgements.

Cards with difference in same direction Cards with difference in opposite direction	Significant, $p < 0·001$.
Card X Card Y	Significant, $p < 0·001$.
Card X Card Z	Significant, $p < 0·001$.
Card Y Card XY	Significant, $p < 0·001$.
Cards P and Q Cards R and H	Significant, $p < 0·001$.
Cards P, Q, S and T Cards R and U	Significant, $p < 0·001$.
Cards P and Q Cards S and T	Significant, $p < 0·01$.

to the contrast of the circles with their own, monocular backgrounds. The backgrounds showed some mixture, tending to alternate between dark and light grey rather than between black and white. With card XY, which combines one card of the X pair with one of the Y pair, one circle composed of a dark grey on black and a white on light grey was seen as stable, silvery and fused, while the other circle, white on black and dark grey on light grey, was described as alternating violently, although the only difference was that these were steps in the opposite directions from their backgrounds on each eye. Here the stereoscopic depth effect was ambiguous, being seen by some of the subjects and not by others.

Thus, with identical figures, one may get either rivalry or fusion. Card Z makes it untenable that some kind of correlation or identity of stimuli falling on corresponding points within the figures gives the cue for fusion, independently of the messages from surrounding areas. It is interesting in this connection that one of the pairs of dark grey and white circles in card Y was sometimes seen as fused and lustrous, while the others were not. This seems to have been a perception of both at once, a kind of superimposition of incompatible stimuli from corresponding points, but this was not adequate for fusion of disparate retinal points and depth perception.

No comparable seeing of both at once, or one through the other, was reported with the red and green circles, although it was sometimes obtained with the blue and yellow ones which were more widely separated in brightness.

Similar results on stereopsis were obtained with the coloured cards in this series: rivalry or suppression of incompatible colours, although present, had little effect on the perception of depth. The larger and more similar in size the differences in intensity in the same direction on the two eyes, the more reliable the impression of depth, whether or not the background or circle colours were rivalling. This impression of depth disappeared when the circle intensities differed in opposite directions from their backgrounds, as in cards U, Y and Z. Where the difference became very small, as in cards R and H, the depth effect was decreased but not altogether destroyed. However, this might simply be an artefact of the experimental cards, which were not ideal for these tests: the circles were stuck on the cards and might have thrown a slight shadow at the edges, which would provide a supplementary, though fragmentary, contour, perhaps sufficient to account for half the subjects sometimes seeing the depth effect. A marked difference in intensities between the circles themselves, when one was relatively close to the background intensity, also decreased the ease with which they gave rise to a perception of depth. Cards S and T were less reliably fused than P and Q, although when the difference from the background was above a minimum for both, as with card X, an equally big difference in intensity between the circles had no detrimental effect.

The colour selection effects described in the first series also occurred here, the circles often appearing patchy and fluctuating, but sometimes all in one colour. This effect occasionally appeared in an even more extreme form with card H, where eight subjects sometimes saw both circles and background in green, or both in red. The colours, in fact identical, did not appear so, for the circles seemed bright and the background duller or muddy. This was probably partly due to some colour mixing of the backgrounds which made them duller in colour, and perhaps partly to simultaneous colour contrast of the circles with their own background. This again would be evidence of an apparently temporarily "suppressed" stimulus having an effect on perception.

Since it has been shown that an impression of stereoscopic depth can be caused, in the absence of other cues, by figures with the requisite position disparity and an intensity difference in the same direction from the background, the question arises whether this is the only stimulus which can be used in this way. It was obviously of interest to try the converse experiment, to see if similarity of colour could outweigh opposite intensity differences and provide an adequate stimulus for stereoscopic fusion and depth.

(3) *Stereopsis from colour cues*

The third section of this experiment was designed to investigate this possibility. Five more cards were prepared, again with different intensity relationships between circles and backgrounds, but this time all the circles were coloured green with the same ink, which had been diluted to give different saturations on white paper, and therefore also different intensities. The details of cards, intensities and the results of the tests on stereopsis and diplopia are given in Table III. Cards L, M_2 and N_1 are similar in their intensity patterns to cards X, Y and Z.

The results with these cards showed the same trends as with the black, grey and white circles of cards X, Y and Z, but significantly more subjects did gain an impression of stereoscopic depth with cards M_1, M_2 and N_1 than with cards Y and Z. This seems to indicate that colour can provide a cue of similarity to give rise to

TABLE III

Card	Circle colours	Log foot lamberts	Back-ground colours	Log foot lamberts	Intensity differences	Depth perceived		Diplopia		No. of S's
						Always	Occa-sionally	Always	Occa-sionally	
L	Dark green Light green (O)	0·85 1·25	White	1·53	−0·68 −0·28	16	0	1	1	16
M_1	Dark green Light green (O)	0·85 1·25	Grey	1·06	−0·21 +0·19	9	4	6	2	16
M_2	Darker green Lighter green (O)	0·62 1·30	Grey	1·06	−0·44 +0·24	2	3	11	3	14
N_1	Medium green (S)	0·95	Black White	0·43 1·53	+0·52 −0·58	4	1	12	4	16
N_2	Very dark green Very light green (S)	0·50 1·37	Black White	0·43 1·53	+0·07 −0·16	5	8	1	0	16

Significance of χ^2 comparisons between cards, on presence or absence of stereoscopic depth perception. Here, as with Table II, cards and cells were grouped where necessary to obtain sufficiently large expected values in each cell.

Cards with difference in same direction / Cards with difference in opposite direction } Significant, $p < 0{\cdot}001$.
Card L / Cards M_1 and M_2 } Significant, $p < 0{\cdot}001$.
Card L / Card N_2 } Significant, $p < 0{\cdot}001$.
Cards Y and Z / Cards N_1, M_1 and M_2 } Significant, $p < 0{\cdot}001$.
Card L / Card N_1 } Significant, $p < 0{\cdot}001$.
Card M_1 / Card M_2 } Significant, $p < 0{\cdot}05$.

fusion and stereopsis, although not a powerful enough one to counteract completely the conflicting information about figure-ground intensity relationships. This information was still affecting the results, as can be seen by comparing L with M, and M_1 with M_2. M_2 is closest to the intensity values for card Y and gave rise to least stereopsis.

Some interesting observations were made on the apparent brightness of the circles. With card L, dark and light green circles on a white card, four out of 16 subjects did not notice there were two greens, a dark and a light; four said there was a sheen or velvety appearance, and eight saw the colours either rivalling, patchy or with some diplopia. With card M_1, the same greens as before on an intermediate grey card, fourteen subjects saw the two colours clearly alternating or patchy, six with diplopia, while only two described them as velvety or with a sheen. With card N_1, where the circles were in fact the same medium green, one on a black and one on a white card, all the subjects saw them as different, some finding it difficult to believe they had not been, when told so afterwards. Twelve always failed to fuse the circles and four sometimes did. The circles on the black ground were seen as much lighter than those on the white,—a difference subjectively as marked as in card M_1. On the other hand, the black and white backgrounds did nearly always fuse to some extent and were seen as a "metallic," "elephant" grey, or sometimes a "luminous" grey "with a sheen." The difference in appearance of the circles seemed to be partly an effect of simultaneous contrast, and the fact that it was so marked even though the black and white of the backgrounds were not separately perceived supports the view suggested by Asher (1950) that the neural mechanisms responsible for contrast effects of this kind must be peripheral to the level of binocular fusion at which the backgrounds are "mixed."

With card N_2 an attempt was made to cancel out this contrast effect to see if rivalry could be eliminated and stereopsis induced even though the intensity differences were still in opposite directions. A lighter green pair of circles was used on the white ground and a darker pair on the black. However, in order to get anywhere near cancelling the strong contrast effects, the intensities of the greens had to be made so widely different that the light green was not much darker than the white and the dark green not much lighter than the black. Under these conditions, depth was usually perceived, but even then the pale green were still seen as slightly darker than the dark green circles. The backgrounds were more or less fused and grey again and the circles were seen as almost the same brightness as the background, or very slightly darker or lighter. The whole picture appeared much more stable and dull than with N_1.

(4) *Contrast and binocular perception of brightness*

One possible explanation of these findings with cards N_1 and N_2 may be that what is coded at a peripheral level is the difference in intensity of the figures from their own monocular backgrounds. Relative excitation or inhibition across the boundaries, or the rates of firing of the "on" and "off" fibres in the retina as the eye moved, would give this relative brightness. The backgrounds are then fused more centrally and the relation "circle on left retina is lighter than the background by amount x" is interpreted perceptually in relation to the fused background, so that the circle is seen as lighter than this fused field by amount x. This might appear to be another way of describing simultaneous contrast, but it defines a little more fully what might be happening in a situation with conflicting binocular information, and it also leads to some different predictions: if a dark grey circle were presented on a black ground to one eye, and a white one on a light grey ground to the other eye, the contrast should be in the same direction for both eyes, if, as seems likely, it is a peripheral process; but on the hypothesis put forward here, the white should appear grey, lighter than the fused background, but darker than the monocular white. A card (W) was prepared to test this point. At the same time, an attempt was made to quantify to some extent, with the apparatus immediately available, the changes in brightness when various pairs of the cards already described and a few new ones were seen combined in the stereoscope. The brightnesses were compared with those seen when the two cards in the stereoscope were identical for the two eyes. This was to see whether the changes in apparent brightness could all be attributed to the normal simultaneous contrast effects.

A series of greens and a series of greys of eight different intensities from 1·45 to 0·50 log foot lamberts were prepared, again using the same inks at different saturations. Rectangles of these 6 cm. by 4 cm. were stuck in ascending order from left to right on to a white card. Subjects were shown the cards in the stereoscope and asked to match the brightnesses of both circles and backgrounds, as seen under those conditions, with the nearest sample rectangle on the matching card, or to say between which two rectangles it lay. A check was kept that they spent approximately the same time looking before taking their decision (about 20 to 30 sec.), to avoid too much variability due to changes in contrast with time. The matching sample card was outside the stereoscope in as near as possible to the same illumination. The subjects were also shown each card with an identical one to the two eyes, and asked again to make a successive comparison and match between the sample rectangles and the circles as they appeared on their own backgrounds seen with both eyes. Each card in all the combinations used was shown once to the right and once to the left eye. The details of the cards and the mean intensities of the samples with which they were matched are given in Table IV.

Clearly there is a considerable difference, particularly with cards N_2, W and V_1, V_2 and V_3, between normal simultaneous contrast with the same picture to both eyes and the brightness matches made with the non-identical pairs presented inside the stereoscope. Seventeen out of 20 of the differences between brightness matches

TABLE IV

APPARENT BRIGHTNESS IN THE STEREOSCOPE AND WITH NORMAL SIMULTANEOUS CONTRAST

Card	*Circle intensities*	*Background intensities*	*Contrast steps in intensity*	*Normal match*	*Stereoscopic match*	*Significance of "t"*	*Stereoscopic background match*	*No. of S's*	*Predicted stereoscopic match*
Z	Grey, 0·80 Grey, 0·80	Black, 0·43 White, 1·53	+0·52 −0·86	0·95 0·67	1·15 0·66	Sig. $p < 0\cdot01$ Not sig. $p < 0\cdot9$	0·74– 1·20	12	1·04 0·58
N_1	Green, 0·95 Green, 0·95	Black, 0·43 White, 1·53	+0·52 −0·64	0·95 0·89	1·15 0·73	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·71– 1·22	12	1·08 0·75
N_2	Green, 0·50 Green, 1·37	Black, 0·43 White, 1·53	+0·22 −0·17	0·65 1·36	1·09 1·05	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·95– 1·12	12	1·09 0·98
N_1 N_2	Green, 0·95 Green, 1·37	Black, 0·43 White, 1·53	+0·52 −0·17	0·95 1·36	1·17 0·94	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·77– 1·13	12	1·22 1·06
N_1 N_2	Green, 0·95 Green, 0·50	White, 1·53 Black, 0·43	−0·64 +0·22	0·89 0·65	0·79 0·99	Sig. $p < 0\cdot01$ Sig. $p < 0\cdot001$	1·20– 0·86	12	0·67 0·93
W	Grey, 0·53 White, 1·53	Black, 0·43 Grey, 1·21	+0·31 +0·30	0·74 1·51	1·02 1·26	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·91– 1·00	12	1·17 1·21
V_1	Green, 0·95 White, 1·53	Black, 0·43 Grey, 1·21	+0·22 +0·30	0·65 1·51	1·12 1·29	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·81– 1·00	12	1·28 1·29
V_2	Green, 0·50 White, 1·53	Black, 0·43 Grey, 1·21	+0·22 +0·30	0·65 1·51	0·93 1·28	Sig. $p < 0\cdot001$ Sig. $p < 0\cdot001$	0·89– 0·98	12	1·08 1·18
V_3	Green, 1·37 White, 1·53	Black, 0·43 Grey, 1·21	+0·97 +0·30	1·40 1·51	1·43 1·36	Not sig. $p < 0\cdot2$ Sig. $p < 0\cdot001$	0·74– 0·87	11	1·58 1·31
J	Green, 0·50 White, 1·53	Black, 0·43 Black, 0·43	+0·22 +1·17	0·65 1·60	0·70 1·51	Not sig. $p < 0\cdot4$ Sig. $p < 0\cdot02$	0·43	11	0·89 1·36

The column giving the significance of Student's "t" gives the significance of the difference between the apparent brightness with the two eyes seeing identical cards ("Normal match") and the apparent brightness when the cards were seen in other combinations in the stereoscope ("Stereoscopic match").

Significance of some further differences between apparent brightnesses, all seen in the stereoscope.
N_1 on white card with N_2 on black > N_1 on white with N_2 on black, significant, $p < 0\cdot05$.
N_2 on black card with N_2 on white > N_2 on black with N_1 on white, significant, $p < 0\cdot001$.
N_1 on black card with N_2 on white > N_1 on black with N_1 on white, significant, $p < 0\cdot05$.
N_2 on white card with N_2 on black > N_2 on white with N_1 on black, significant, $p < 0\cdot05$. (Change in reverse direction.)
Brightness change in V_2, white circle on black > Brightness change in V_1, white circle on grey, significant, $p < 0\cdot001$.

The standard deviations of the "stereoscopic matches" ranged between 0·03 and 0·18, with a median of 0·08, and the standard errors ranged from 0·01 to 0·05 with a median of 0·02. The standard deviations of the "normal matches" ranged from 0·04 to 0·10, with a median of 0·06, and the standard errors ranged from 0·01 to 0·03, with a median of 0·02.

under the two conditions were statistically significant, at least at the $p < 0\cdot05$ level. With card N_2 there was little or no change due to normal brightness contrast, but in the stereoscopic combination, the change was so marked that the pale green on the white was matched with a slightly darker green than the match made to the dark green circle on the black card. The match for the "fused" backgrounds was often found difficult, for as soon as it became the object of attention, it tended to fluctuate or vary between a darker and a lighter grey. The difficulty was particularly marked with cards where the differences in intensity between background and figures were largest. Where it was found impossible to choose a single grey to match it, the greys between which the ground was rivalling were given instead.

With cards W, V_1, V_2 and V_3 it is no longer possible to hold simultaneous contrast responsible for the darkening of the white circles, for both backgrounds were darker than the circles. With card W the circles were seen as two similar shades of grey (by three subjects as the same grey for the first minute or two), although the whiter pair were always later seen as a little lighter than the dark grey ones. Could this darkening have been due simply to mixing of the circle brightnesses, rather than a relative figure ground effect of the sort hypothesized? With cards V_1 and V_2 there did not appear to be any trace of green in the darkened white, apart from the usual patchy, rivalling effect. Also, with card V_3 the darkening of the white actually made it significantly darker ($p < 0\cdot01$) than the green on the other eye, so that some process other than averaging the figure brightnesses seems to have taken place. However, mixture or "averaging" of the circle intensities was playing some part, as could be seen by the different matches made to N_1 and N_2 when they were differently paired, keeping the same backgrounds. Card J was made to test how much of the darkening could be attributed to mixing, for here the backgrounds were identical, both black, and only the circles were different. A slight darkening of the white did take place, but significantly less ($p < 0{\cdot}001$) than with card V_2 where the white circle was paired with the same green on black as card J, but was itself on a grey background.

There seemed in fact to be some compromise between this "averaging" of the circle brightnesses, and the system suggested earlier, where the difference in intensity from the monocular background was interpreted as referring to the binocular "fused" background. Taking this "compromise" hypothesis at its crudest and simplest for the moment, one can attempt to see if it would lead to predictions in the correct direction for all the matches made. The predicted values given in the table for the apparent brightness of the circles are calculated by the following somewhat arbitrary method. The two factors determining the perceived brightness were taken to be (1) the average of the brightness effectively reported by each eye (shown by the matches made with identical cards seen with both eyes). (2) The brightness given by relating the apparent contrast difference in intensity at the eye, not to the background at that eye, but to the background resulting from fusion. These two contributions were averaged to predict the subjects' matches in the binocular situation. The formula was: $\dfrac{\bar{B} + (MC - MB) + \bar{C}}{2}$, where $\bar{B}$ is the fused background $(MC - MB)$, the difference in intensity between the monocular background and the apparent brightness of the figure presented to that eye (with the normal contrast effects as determined with identical figures to the two eyes), and $\bar{C}$ is the mean of the apparent intensities of the two circles (again with the normal simultaneous contrast effect). For cards where the backgrounds were only partly fused, $\bar{B}$ was taken as the semi-fused light grey for the circle on white, and the dark grey for the circle originally on black.

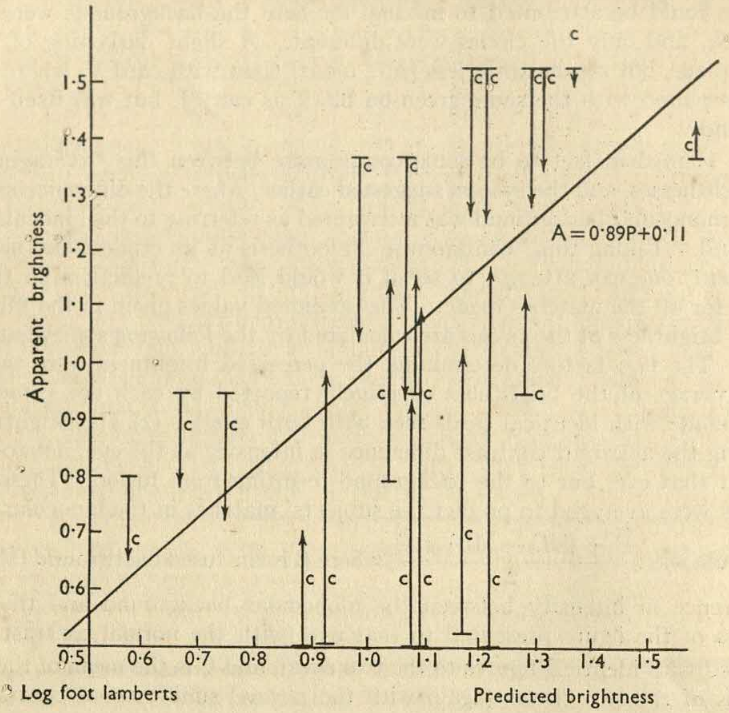
The results fit these predictions well in some cases, and all cases show a change from the monocular intensity in the expected direction; the product-moment correlation between predicted and actual values gives $r = 0{\cdot}90$ (significant, $p < 0{\cdot}001$). The graph shows the actual values plotted against the predicted ones. The monocular intensities are given by the bars (value on the y axis); the arrows show the extent of the change in apparent brightness in the stereoscope, and the letter "c" shows the apparent brightness with the same card to both eyes, that is the normal simultaneous contrast effect. The regression line of the apparent brightness in the stereoscope on the predicted brightness is also given; it is quite close to the predicted 45° relationship, but a little less steep ($A = 0{\cdot}89\ P + 0{\cdot}11$). This seems to be because subjects' judgements of brightnesses at both extremes of the scale were closer to the mean than was predicted.

The product-moment correlation between the monocular figure-ground contrast differences and the apparent brightnesses of the circles in the stereoscope was also high (0·85) (significant, $p < 0{\cdot}001$), while the correlation between these stereoscopically perceived brightnesses and the mean of the circle brightnesses with normal contrast effects was 0·71, the latter being significant at the 0·01 level.

Since the hypothesis was left in a crude form, simply as one suggestion, no more weight will be attached to this point than to say that the fit appears encouraging.

It would, of course, need more quantitative refinement than was possible with the simple apparatus available.

FIGURE 2



Regression of apparent brightness on predicted brightness. The bars show the monocular intensity of the figure, the arrows the extent of the change to the apparent brightness, and the letter "C" indicates the normal simultaneous brightness.

General Discussion

It would be rash to speculate in much detail on the neurophysiology behind these findings. Several points do seem to emerge from the experiment, however. It seems clear that colour information can be treated separately from the positions and shapes. If one supposes that the suppression in rivalry occurs somewhere at the periphery, then from the suppressed areas there must be many input fibres which are left uninhibited. It seems unlikely that Bárány's finding on the loss of the pupillary reflex would hold in this situation. There seem to be at least three types of perception: a random patchy network, suppression of one colour as a whole, or suppression of signals from one retina as a whole. The first could be explained by competition between "corresponding fibres" for individual common pathways at some more central level, giving a random distribution of the two colours. De Valois (1960) has identified cortical cells in the monkey which respond selectively to a red light presented to either of the two eyes.

The other two types of perception perhaps imply some central or centrifugal inhibitory mechanism of a more systematic type. In all the cards with compatible intensity relationships, there was rivalry between differing colour information only.

This may give additional support to the view that colour information is carried by different neurons from those which respond to intensities of light (Granit's modulators (1955) perhaps, leaving the dominators to carry the information on intensities of the stimuli). Neurophysiological evidence for the relative independence of these two types of information is also given by De Valois who found cells in the dorsal layers of the monkey's lateral geniculate nucleus which responded specifically to different wavelengths and were not affected by the state of adaptation. Also interesting, from the point of view of colour rivalry, is his finding of a second colour receptive system in the intermediate layers where the same cell responded with the "on" responses to one colour and with "off" responses to its complementary colour.

The perception, often reported here, of a card seen entirely in one colour, although this involved areas on both eyes, could be explained as a central or centrifugal suppression of these specialized cells carrying colour information.

The perception of stereoscopic depth appears to be dependent not only on slight displacements of identical contours falling within the fusional areas, as is generally accepted, but also on the particular intensity relation between figure and background. This complements Fry and Bartley's finding that binocular brightness rivalry was dependent on the intensity level of the field lying between the intensities of the rivalling patches. Grindley and Whittle (personal communication) are also investigating these aspects of binocular rivalry at present and some of their results are similar. As has been said earlier, it is tempting to suggest that "fusion" and stereopsis are dependent on similar excitatory, or alternatively inhibitory, messages from the two eyes. One plausible system might be the "on" and "off" signals, which have been recorded both peripherally (Kuffler, 1953) and centrally (Hubel and Wiesel, 1959). These give a different response to changes in intensity in opposite directions from the resting level, and together with eye movements are often thought to be useful in giving predominance to contours or discontinuities in distribution of light. They also convert intensities into "figure-ground" relationships rather than absolute levels; this appears to fit the requirements here found experimentally for "fusion" and depth perception. These signals may, then, be compared from each eye, and if both are "on" or both "off" responses, disregarding the absolute frequency, (apparently over a wide range: 0·9 log foot lamberts in card W) they are combined, and can give rise to perception of depth, if the requisite horizontal disparity is also present. Granit (1955) cites considerable evidence for the antagonistic, mutually exclusive nature of the "on" and "off" responses, at least at the retinal level. A binocular neurophysiological interaction which may be of a similar type to this has been shown to occur by Hubel and Wiesel (1959). They recorded from "binocular" units in the cat's striate cortex and found interaction and summation of responses between excitatory and inhibitory regions of the two homologous receptive fields. Of interest here is their observation that an inhibitory signal from one eye could cancel an excitatory one from the other in a single cortical cell. They also found many cells responding only to one eye, and these might supply the signals which are seen as rivalling; possibly, if there were only binocular cells, incompatible signals would "cancel" each other and never be seen.

The explanation in terms of "on" and "off" responses to increases or decreases of light stimulation might be dependent on the occurrence of co-ordinated eye movements in the two eyes to maintain the initial stereopsis. A preliminary test with just three subjects in fact suggested that when a central point was fixated the depth effect disappeared, the circles were seen in one plane and tended to become unfused. This may be some additional support for the "on" and "off" signals hypothesis.

It would, however, also be possible to have a similar explanation for a stationary field. Hubel and Wiesel showed that the retinal receptive fields of cortical cells consisted of elongated central areas, about one degree wide and four long, surrounded by areas of opposite excitatory or inhibitory effect. If, for example, part of a figure brighter than the background on one eye lay on a central excitatory field, the balance of excitation would be higher than the inhibition for that eye. But on the corresponding receptive field of the other eye, the background would be brighter than the figure, and would be falling on the flanking inhibitory areas; this might give rise to the antagonistic or cancelling effects found by Hubel and Wiesel. On this hypothesis, one might explain the size of the fusional areas for stereopsis by the size of the central areas in the receptive fields of the binocular cortical cells (both being about one degree wide, although the latter were, of course, measured for the cat's retina).

Along similar lines one might explain why the backgrounds appeared to "fuse" or "average" in brightness more than the circles: where there are no changes or discontinuities in intensity, there will only be messages from cells which respond more continuously to light, no signals from incompatible fibres, and therefore perhaps the possibility of taking some kind of average reading of the intensities at the two eyes and seeing this as a more or less "fused" area, as described for uniform fields by Fechner's paradox.

The compatibility of the excitatory or inhibitory messages from each eye does not, however, appear to be the only one for stereopsis, although it plays a large part. Similarity of colour can to some extent compensate for incompatibilities in brightness information to give a more or less stable impression of depth. Coincidence in the firing of cells specific to certain wavelengths appears to provide a supplementary cue, its lesser effectiveness being consistent with the small number of "modulators" as opposed to "dominators."

To return to the apparent brightness and contrast effects noticed independently of the depth effects: the findings with cards Z and N_1 (that rivalry was induced between identical circles because of their different monocular backgrounds) could be explained by saying that simultaneous contrast, excitation or inhibition, occurs peripherally to the binocular system mediating stereopsis. However, when one compares in more detail the matches made with different cards at each eye and the normal contrast matches made with the same card seen by each eye (particularly with cards N_2, V and W) some striking differences emerge which suggest that one needs to modify the explanation based simply on monocular contrast. One might suppose that, while the "on" and "off" fibres (or relative excitation and inhibition generated at the figure-boundaries) from the two eyes are not compatible and are never combined in the binocular image, the fibres responding continuously to illumination might combine, as Fechner's paradox suggests, to give an average of the two monocular circle brightnesses, as well as of the backgrounds. The perceived brightness of the figures might be some compromise between (1) the average of the two brightnesses seen with simultaneous contrast effects from the background, which would be the contribution of the fibres giving a continued response to illumination and (2) the mean fused background plus or minus the difference of the figures from their monocular background. The latter would be the result of taking relative brightness from the rate of firing of the "on" and "off" fibres and perceiving it in relation to the "fused" averaged background. As was shown in Table III, the results are sufficiently similar to the predictions, even on this somewhat arbitrary calculation, to suggest that the underlying processes may be of the general type suggested.

I should like to thank Professor R. C. Oldfield for helpful criticism, my husband, Dr. M. Treisman, for his advice and encouragement, Mr. P. Whittle for an interesting discussion, and all the volunteer subjects who took part in the experiment. The work was supported by a grant from the Medical Research Council.

REFERENCES

ASHER, H. (1950). Contrast in eye and brain. *Brit. J. Psychol.*, **40**, 187–94.

BAGBY, J. W. (1957). A cross-cultural study of perceptual predominance in binocular rivalry. *J. abnorm. soc. Psychol.*, **54**, 331–4.

BÁRÁNY, E. H., and HALLDEN, U. (1948). Phasic inhibition of the light reflex of the pupils during retinal rivalry. *J. Neurophysiol.*, **11**, 25–30.

BARTLEY, S. H. (1941). *Vision: A Study of Its Basis.* New York.

BREESE, B. B. (1909). Binocular rivalry. *Psychol. Rev.*, **16**, 410–5.

BURIAN, H. (1936). Studien über zweiäugiges Tiefensehen bei örtlicher Abblendung. *V. Graefes. Arch. Ophthal.*, **136**, 172–214.

CREED, R. S. (1935). Observations on binocular fusion and rivalry. *J. Physiol.*, **84**, 381–92.

DE VALOIS, R. L. (1960). In Y. Galifret (Ed.) *Mechanisms of Colour Discrimination.* London.

DITCHBURN, R. W. (1955). Eye movements in relation to retinal action. *Opt. Acta*, **1**, 171–6.

ENGEL, E. (1956). The role of content in binocular resolution. *Amer. J. Psychol.*, **69**, 87.

FRY, G. A., and BARTLEY, S. H. (1933). The brilliance of an object seen binocularly. *Amer. J. Ophthal.*, **16**, 687–93.

GRANIT, R. (1955). *Receptors and Sensory Perception.* Yale.

HANES, R. M. (1949). A scale of subjective brightness. *J. exp. Psychol.*, **39**, 438–52.

HUBEL, D. H., and WIESEL, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.*, **148**, 574–91.

KUFFLER, S. W. (1953). Discharge patterns and the functional organization of the mammalian retina. *J. Neurophysiol.*, **16**, 37–68.

OGLE, K. N. (1950). *Researches in Binocular Vision.* London.

RØNNE, G. (1956). The physiological basis of sensory fusion. *Acta Ophthal.*, **34**, 1–26.

A STATISTICAL TEST OF A PREDICTION FROM INFORMATION THEORY IN A CARD-SORTING SITUATION

BY

D. R. J. LAMING

From the Department of Experimental Psychology, University of Cambridge

Previous work with cards left the joint effects of varying entropy and discriminability on choice-reaction times somewhat in doubt. A card-sorting situation is here constructed in which entropy and discriminability can be varied independently. Two experiments are described with naive subjects and two different treatments of the practice effect. The effect of varying entropy with discriminability held constant is shown to be significantly non-linear, contrary to the prediction from information theory. Finally a statistical model is proposed which shows at least qualitative agreement with the results of these experiments.

INTRODUCTION

Previous experiments on choice-reaction times have mostly had either a constant difficulty of discrimination or a constant entropy of the choice situation. Hick (1952) showed that equi-probable multi-choice-reaction times are approximated by the formula $t = a \log (n + 1)$, ($n = 2 \ldots 10$). He concluded that the amount of information extracted from the situation is proportional to the time taken to extract it, on the average, and found this to be true also when a large number of errors was deliberately made. Hyman (1953) obtained results fitting the formula $a + b \sum_{i} p_i \log p_i$. He also found that the mean response times to several alternative signals were different in conditions of frequency unbalance, but not so different as a consideration of their respective entropies would lead one to expect. Crossman (1953) also investigated conditions of frequency unbalance, and found some effect of discriminability on reaction time. Crossman (1955) examined the reaction time for several different levels of discriminability, and found reaction times approximately linear with $C (S_1S_2) = 1/[\log x_1 - \log x_2]$, which he called a "confusion function." When this confusion function was generalized to more than two choices, it seemed to give a complete explanation of reaction time, regardless of the increase in entropy. But, apart from this one experiment, the effects upon choice-reaction times of entropy and discriminability have usually been studied separately.

If we adopt Wiener's (1948) formulation of the amount of information required to make a discrimination, choice-reaction time should be a sum of separate delays due to discriminability on the one hand and frequency of occurrence on the other. We should expect, then, that pairs of signals with different discriminabilities would give a family of parallel straight lines on a time-entropy plot. Crossman's confusion function may give a similar result, but his final experiment (1955), provided his generalization of the confusion function was valid, suggests some form of interaction. The following experiments set out to vary both discriminability and entropy systematically, to see whether their respective effects upon choice-reaction time summate simply or interact.

Two arbitrary features of these experiments should perhaps be justified; the use of naive subjects, and the use of a two-choice situation.

The subjects used in these experiments were given only a very small amount of practice, designed to accustom them to the experimental situation—to enable them, for example, to determine beforehand the most convenient way to hold the cards for sorting. In this sense we call these subjects "naive," to distinguish them from subjects who are given practice at the experimental task in order to improve their performance to some stable level of efficiency. These latter we call "practised."

Naive subjects were used for these experiments because it is held that the mode of functioning of the naive subject is of greater importance than that of the highly practised one. Hick and Crossman showed that their particular subjects were capable, given sufficient training, of acting as an information channel of constant capacity, but their results can claim no greater generality than this. There is no way of showing that the level of practice at which Hick and Crossman made their measurements was the most efficient obtainable and not merely a plateau. Mowbray and Rhoades (1959) have found that with long practice the information slope tends to disappear. If, then, we practice our subjects, we must justify the particular amount of practice we give them, else our results have no claim to universality. If we use unpractised subjects, our results may be valid for all unpractised subjects, which is nearly everybody in the world.

It is possible that with practice the subject may develop a different mode of functioning. If this is so, it needs to be shown; but the functioning of the naive subject will still need to be studied.

A two-choice situation was chosen in case the result of Crossman's final experiment (1955) was due to his generalization of his confusion function to more than two dimensions not being valid. We have achieved in these experiments a variation of entropy with constant discriminability, whatever form the confusion function might take, provided it does not involve probabilities of occurrence. A two-choice situation also has the virtue of keeping the number of independent variables to a minimum, which is a vital consideration in validating any mathematical model. Although a number of models that have seemed to fit a two-choice situation have proved impossible of generalization to three or more choices, it is nevertheless necessary to begin with the study of two-choice situations, or we are faced with mathematics of appalling complexity.

METHOD

In these experiments we used only a two-choice situation. Discriminability was varied by changing the physical magnitudes of the discriminanda, and entropy by changing their relative probabilities of occurrence. These two factors could be varied quite independently. We shall use "ratio" to signify the ratio of the physical sizes of the discriminanda, "frequency" to signify their relative probabilities of occurrence, expressed as the number of occurrences of each signal in a series of 48, and "frequency scale" to signify a scale linear in the probability of occurrence of a given signal.

Preliminary experiments in the present series suggested that choice-reaction time is linear, not with entropy, but with frequency, that the gradient of the time-frequency line increases with the ratio approximately according to the function $(\text{ratio} - 1)^{-\frac{1}{2}}$ or $(\log \text{ratio})^{-\frac{1}{2}}$, and that the absolute reaction time for a particular subject varies appreciably from session to session. More errors were made in these preliminary experiments on the infrequent than on the frequent signal. Accordingly, in the main experiments the reaction time was determined for three different frequencies, so chosen as to provide a sensitive test between the "time linear with entropy" hypothesis and "time linear with frequency." A test between the functions $(\text{ratio} - 1)^{-\frac{1}{2}}$ and $(\log \text{ratio})^{-\frac{1}{2}}$ was carried out in the same way, and the measurements from each subject, in order to be comparable, were taken all in the same session.

Within an experimental session we found a progressive decrease in reaction times due to practice. Hick and Crossman used only a small number of practised subjects;

their measurements were made when their subjects had reached a state of level performance. But since reaction times are also affected by motivational factors and boredom, such uniformity of performance is somewhat artificial. Further, high degrees of practice are not practicable with a statistical test of the hypotheses in question against the variance between subjects, since this requires many subjects. There are two other ways of meeting practice effects; either we have three random groups, one for each frequency (Experiment I), or we have six random groups, each group tested at all three frequencies, but in a different order (Experiment II).

Choice-reaction times were obtained by measuring the time taken to sort a pack of 48 cards. The cards were blank playing-cards with black stripes painted on them in indian ink. A template was first prepared by cutting a hole the exact size of the stripe required in the centre of a card. The template was used to outline the stripe in pencil on each card. Because of the properties of surface tension relative to the shiny surface of the cards and the pencilled outline, it was possible to fill in the outline to a tolerance of about 0·01 in. The stripes were 0·26 in. wide and 2·83 in., 1·68 in., 1·41 in. or 0·35 in. long, and were staggered randomly along the centre-line of the card. The cards were randomized for sorting by shuffling in a uniform manner—three "riffles" and a "cut," always done by the experimenter.

The subjects were volunteer university students. They were seated at a table, sorting the cards into two piles. Each pack began with a blank card; the subject began to sort when he felt ready and the time from the blank card to the 48th card, an exact 48 sorting movements, was measured with a stop watch to the nearest tenth of a second. These measurements were affected a little by the experimenter's reaction time, but the effect produced was only a small and constant increase in the means and variances of those distributions from which the measurements were a sample. Duplicate packs were used, so that while one was being sorted, the experimenter could check and shuffle the other. Subjects were told what errors they had made one pack in arrears, but never what time they had taken to sort.

EXPERIMENT I

Three random groups (A, B and C) of 12 subjects each were used for this experiment, and packs with three different ratios (1·19, 1·68, 8·00) made up of the pairs of lengths (1·68 in., 1·41 in.), (2·83 in., 1·68 in.), (2·83 in., 0·35 in.). Half of each group worked with a majority of long stripes, half with a majority of short, in the conditions of frequency unbalance, and we shall denote these half-groups by A_l, A_s, etc. Each subject did one session of about 75 min. The session began with a standard amount of practice (two 82-card packs, frequency 41 : 41, ratios 1·09 and 1·19, sorted in the subject's own time), and continued with one practice run and six timed runs on each of seven packs. The specifications of the particular packs sorted by each half-group are given below in Table I. The frequencies are shown by the symbol $x : y$, where x = number of cards with long stripes, y = number of cards with short stripes in the pack.

TABLE I

SPECIFICATIONS OF THE PACKS SORTED BY EACH HALF GROUP IN EXPERIMENT I

Half group	*Pack 1 ratio 1·19 face up*	*Packs 2 to 4 1·19, 1·68, 8·00 face down*	*Packs 5 to 7 1·19, 1·68, 8·00 face down*
A_l	24 : 24	24 : 24	36 : 12
A_s	24 : 24	24 : 24	12 : 36
B_l	30 : 18	30 : 18	18 : 30
B_s	18 : 30	18 : 30	30 : 18
C_l	36 : 12	36 : 12	24 : 24
C_s	12 : 36	12 : 36	24 : 24

The six subjects within each half-group each sorted the three ratios in a different order and kept to the same individual order for packs 2 to 4 and for packs 5 to 7. There was a short break in the session between packs 4 and 5.

The 36 subjects used had no previous experience of this sorting task. Sample stimuli were displayed for the 1·19 ratio in case the subject needed to remind himself of the two lengths of stripe in the pack. The subjects were not obliged to put their piles in particular places nor on particular sides of the table. They were told that it was better to go too slow than too fast, in order to avoid errors. They were asked not to correct their mistakes, but nonetheless some of them did, explaining that it was an immediate reaction. The incidence of these corrections was small and randomly distributed among the subjects. Hence these corrections caused only a slight increase of experimental variance and did not bias the outcome, and so no special allowance for these cases was required.

Results

Table II shows the mean times in sec./pack for each condition in Experiment I. The means are averaged over all the subjects timed under each particular condition, and the variances quoted are those obtained between the mean times for the individual subjects.

TABLE II

MEANS IN SEC./PACK AND VARIANCES FOR EACH FREQUENCY-RATIO CONDITION IN EXPERIMENT I

Frequency	*Entropy*	*Pack 1*	*Packs 2 to 4*			*Packs 5 to 7*		
		1·19 face up	*1·19*	*1·68*	*8·00*	*1·19*	*1·68*	*8·00*
24:24 (means of 12 subjects)	1·00	30·40 18·80	52·75 44·80	41·92 15·29	39·65 15·44	46·22 17·65	38·53 5·02	36·99 4·96
30:18 (means of six subjects)	0·954	31·87 41·36	53·51 84·66	44·36 56·68	43·00 50·94	49·70 159·48	40·62 71·42	38·29 42·77
18:30 (means of six subjects)	0·954	31·11 127·80	48·57 132·38	41·82 70·87	38·89 77·51	50·66 57·53	42·16 21·72	41·01 43·62
36:12 (means of six subjects)	0·812	28·98 41·66	47·87 10·16	39·07 17·96	36·66 2·47	47·37 18·73	36·96 7·51	35·95 4·20
12:36 (means of six subjects)	0·812	26·20 4·40	43·66 4·49	37·88 0·83	35·74 2·89	47·50 54·70	38·63 45·90	35·95 43·03

(i) *A majority of long stripes compared with a majority of short*

The subjects in group B were all timed with packs of both frequency 30:18 and 18:30. The mean time for the frequency 30:18 (44·91 sec.) was significantly greater than for 18:30 (43·85 sec.) at the 0·001 level in an Analysis of Variance. Hence the times for packs with majorities of long and short signals should be analysed separately.

(ii) *The relation between time and frequency*

The frequencies 24:24, 18:30 and 12:36 have entropies 1·000, 0·954, 0·812 respectively. Hence if t_A, t_B, t_C are their respective mean times, we should find $3t_A - 4t_B + t_C \simeq 0$ if the information theory hypothesis is correct. The same relation should hold if we consider the frequencies 24:24, 30:18 and 36:12. On the other hand, the linear-with-frequency hypothesis would predict $t_A - 2t_B + t_C = 0$.

To test these hypotheses we make a comparison within each column of Table II, keeping the analysis of the time for majorities of long signals separate from that for majorities of short. But if we compare the variances within any of the columns of Table II we find that they differ from each other significantly at levels as high as 0·001. Accordingly an Analysis of Variance is not possible with these data, because the sample sizes are too small, and we must test the above hypotheses by the formulae:

$$t = (3t_A - 4t_B + t_C) \Big/ \sqrt{\left(\frac{9\sigma_A^2}{12} + \frac{16\sigma_B^2}{6} + \frac{\sigma_C^2}{6}\right)}$$

$$t = (t_A - 2t_B + t_C) \Big/ \sqrt{\left(\frac{\sigma_A^2}{12} + \frac{4\sigma_B^2}{6} + \frac{\sigma_C^2}{6}\right)}$$

where t has six degrees of freedom, and σ_A, σ_B, σ_C are the variances corresponding to t_A, t_B, t_C respectively. These formulae do not give any significant differences from zero, even when we combine "packs 2 to 4" with "packs 5 to 7."

Table II suggests that there is a rise in time from 24:24 to 30:18. But the difference between the means is certainly not significant: it probably arises because B_l, as a group, was slow, and would not have appeared if the groups had been sufficiently large. The results under this heading are strikingly insensitive compared with those under the previous heading, because in this case it was not possible to remove the variance due to the large inter-subject differences in mean reaction time.

(iii) *Change of the time-frequency relation with varying ratio*

We shall use the difference between the mean times for the same subject at frequencies 24:24 and 12:36 (or 36:12) as a measure of the gradient of the time-frequency curve. In an Analysis of Variance we find no significant difference between the gradients, measured in this way, for the three ratios used. This experiment therefore fails to distinguish between the two hypothetical functions $(\text{ratio} - 1)^{-\frac{1}{2}}$ and $(\log \text{ratio})^{-\frac{1}{2}}$.

(iv) *Unbalance of errors*

In conditions of frequency unbalance more errors are made on the infrequent signal, even when we disregard the difference in number of errors possible. This difference is significant at 0·001 on a χ^2 test for groups B and C. This unbalance of errors seems to become more extreme with unbalance of frequency, but this tendency is not quite significant at the 0·05 level.

EXPERIMENT II

Experiment I failed to yield a decisive answer on the relation between time and frequency because the size of the groups was too small relative to the variance between subjects. If on the other hand we used the variance within subjects to test our hypotheses, we might obtain spurious results arising from one group's being slower than another. For example, if we compare the mean times for the frequencies 24:24 and 18:30 in Table II, we find that for the times obtained from groups B_s, A_l and A_s (packs 2 to 4), frequency 18:30 is faster than 24:24, whereas for the times from groups B_l, C_l, C_s (packs 5 to 7), the reverse is true; both these differences are significant when tested against within-subject variance. It is important therefore to use the variance between subjects to test the hypotheses here. We can reduce the size of the group required to obtain a sensitive outcome by having every subject timed under every condition. The times can then be adjusted for each subject's "movement time" in an Analysis of Variance. This procedure was followed in Experiment II.

Only one ratio (1·68) made up of the pair of lengths (2·83 in., 1·68 in.) and three frequencies (24:24, 18:30 and 12:36), which we shall designate A, B and C, respectively, were used in this experiment, so that in conditions of frequency unbalance there was always a majority of short stripes. There were six random groups of six subjects each, one group to each possible order of the three frequencies. Each experimental session began with one practice run with that pack on which the subject in question would first be timed; then four times were recorded at each frequency in the order appropriate to the subject's order group; after a break of 2 min. or so another four times were recorded at each frequency, but in the reverse order. Each subject did one such session, which lasted about 30 min. The reversal of frequency order in the second half of the session reduced the effect of practice during the session on mean sorting times. Sample stimuli were displayed with the larger stripe always to the subjects' right.

Half of the subjects in each group had had at least 1 hour's previous experience at this task, the others were naive. We shall designate these two groups of subjects as the Practice and the Non-Practice groups respectively. They were all told that the experiment was designed to investigate small differences in the mean times under the three frequency conditions, and were instructed to go as fast as they could throughout the session. But they were also warned not to make too many mistakes—not more than one or two per pack—since this would make their times unreliable. They were explicitly instructed not to correct any mistakes they made, although one or two of them still tried to do so. They were informed of their errors, but not their times, as in Experiment I.

Finally the orders of long and short stripes were noted in eight sample shufflings for each frequency with the pack used in Experiment II. These were tested for randomness by a one-sample runs test. At the 24:24 and 18:30 frequencies there was no significant departure from randomness; at 12:36 there was a small but highly significant departure in the direction of too great a length of run of similar cards. Any effect which this lack of randomness had on the times recorded would have decreased the departure from the information theory prediction.

Results

TABLE III

Mean Time in Sec./Pack in Experiment II

Frequency	24:24	18:30	12:36
Entropy	1·000	0·954	0·812
Practice group	38·04	36·94	35·79
Non-practice group	39·74	39·20	37·67

(i) *The relation between time and frequency*

As in Experiment I we have to test the hypotheses $3t_A - 4t_B + t_C = 0$ and $t_A - 2t_B + t_C = 0$. We can calculate for each subject a departure from each of these two hypotheses; these departures will be approximately normally distributed.

The mean departure from the "time linear with frequency" hypothesis ($t_A - 2t_B + t_C = 0$) is significantly different from 0 at the 0·05 level for all subjects together and for the non-practice group alone. The mean departure from the prediction from information theory ($3t_A - 4t_B + t_C = 0$) is significantly different from 0 at the 0·05 level for all subjects together and at the 0·01 level for the practice group alone. The difference in mean departure between the practice and non-practice groups is significantly different from 0 at the 0·05 level.

(ii) *Unbalance of errors*

TABLE IV

Error Scores in Experiment II

Frequency	A = 24:24	B = 18:30	C = 12:36
Maximum number of errors possible	3,456, 3,456	2,592, 4,320	1,728, 5,184
Practice group	54, 29	63, 16	69, 18
Non-practice group	105, 54	112, 39	143, 13

where (x, y) indicates x errors on the long stripes, y errors on the short.

There are more long than short errors made at all frequencies; this is significant against individual differences at the 0·0005 level for A and B, and the 0·0001 level for C by a 2-tailed sign test. The unbalance of errors in frequency A may be due to the effect of doing frequency B or C previously. If we compare the first four with the last four trials in the session of groups ABC and ACB we find that the balance of errors with frequency A was initially about even, but that it had changed by the end of the session; this change is significant at the 0·012 level on a 2-tailed sign test. There is no similar effect significant for frequencies B and C.

The balance of errors, as measured by (long errors—short errors) is more extreme for C than for B; this tendency is significant at the 0·005 level on a 1-tailed sign test, and since the overall error level for each individual subject was roughly constant, this is a valid way of testing the hypothesis. The comparison between A and B was not significant.

The practice group made fewer errors than the non-practice group; this was significant at the 0·02 level on a 2-tailed Mann–Witney U-test. Their ratios of long errors : short errors were also different.

If we calculate the prediction of information theory for this situation with the correction for errors given by Shannon and Weaver (1949, p. 36), (Information transmitted $= H(x) - H_y(x)$), we find a very small average departure, provided all the subjects are taken together. But if we take the practice group alone we find a departure significant at the 0·001 level in the direction of a flatter curve on the time-frequency plot. The non-practice group has a departure significant at the 0·05 level, giving a curve more bowed than that which information theory would predict.

DISCUSSION

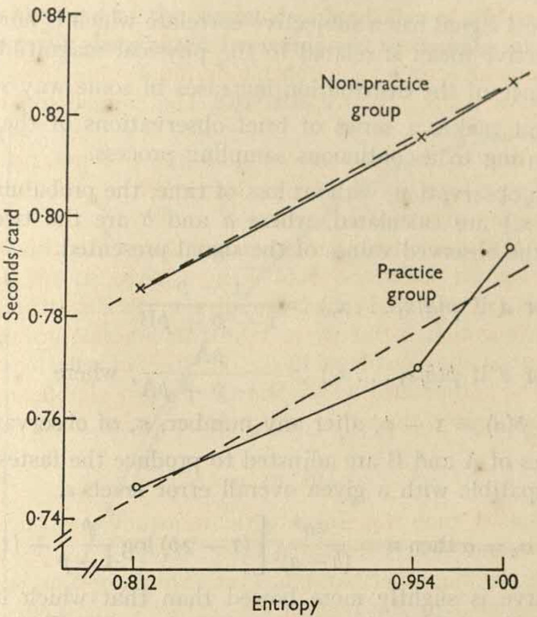
The hypothesis that time was linear with frequency was purely empirical with no theoretical basis and clearly does not agree with the results of Experiment II. Information theory, on the other hand, has hitherto gained support from numerous results. But it has not, to the best of the author's knowledge, been previously subjected to a statistical test against inter-subject variance in such a situation as has been studied here. This is a necessary preliminary to any claim to universal validity; and the results need careful examination.

The time scores from Experiment II are shown graphically in Figure 1, together with the best fits obtainable with a simple information model. The general trend of the variation of time with entropy is close to a straight line, but for the practice group the departure is significant at the 0·01 level. Further, there is a significant difference between the shape of line achieved by the two groups, and it seems likely that the time-frequency relation is not a simple function of time and frequency, but involves other factors as well, error-rate and ratio perhaps. But it may clearly be approximated in most cases by $t \propto -\sum_i p_i \log p_i$.

If we apply Shannon's correction for equivocation in a noisy channel, we find that significant differences still exist, and the agreement when both groups are taken together is almost certainly fortuitous. Information theory, again in the simple form in which it has been applied here, would predict no difference between the mean times for the frequencies 30 : 18 and 18 : 30, and is non-committal on the distribution of errors.

The evidence on the effect of varying ratio on the time-frequency relation is inconclusive. Experiment I did not even establish that the gradient was significantly affected by the ratio, although some increase of gradient with decrease of ratio was certainly suggested. As the ratio is reduced towards 1 it reaches a threshold value, which varies from subject to subject and decreases with practice, and around which the errors rise to a chance level. Near this limit error rate is a much more important factor in determining response time than is frequency. There is probably no discontinuity in the behaviour of the response-time function as the ratio is reduced, and so any complete treatment of this problem will need to relate response-time gradient not only to ratio, but to error rate and the threshold value as well.

FIGURE 1



The time scores from Experiment II with the best-fitting straight lines.

These experiments suggest neither a simple summation nor an interaction of the effects upon choice-reaction time due to discriminability and entropy. Unless the true confusion function involves probabilities of occurrence, we may claim that discriminability was held constant in Experiment II. Nevertheless, significant departures from the formula $t \propto -\sum_{1} p_i \log p_i$ were obtained. Rather than discriminability and entropy having separate effects, which may interact with each other, it seems that we should regard the results obtained hitherto as approximations in different situations to the behaviour of some, as yet unknown, model, which will probably involve other factors besides discriminability and entropy.

Any model successful in this experimental situation needs to explain:—

(i) The time-frequency relation, which seems to involve more factors than just time and frequency, and cannot always be described by $t \propto -\sum_{1} p_i \log p_i$.

(ii) The different times taken to decide on frequent and infrequent signals (Hyman, 1953).

(iii) The increase in time with a majority of long signals as opposed to a majority of short.

(iv) The unbalancing of errors, which increases with the unbalance of frequency.

(v) Certain changes with practice, in that the time-frequency curve is flattened, the gradient is reduced (Mowbray and Rhoades, 1959), the error level is reduced and the unbalance of errors seems to become less extreme.

THEORETICAL CONSIDERATIONS

Most of the above requirements can be met at least qualitatively by a model based on the Sequential Probability Ratio Test developed by Wald (1947). We set up a model with the following postulates:—

(i) The physical signal has a subjective correlate which is normally distributed. The subjective mean is related to the physical magnitude of the signal.

(ii) The variance of the distribution increases in some way with the mean.

(iii) The subject makes a series of brief observations of the signal presented, approximating to a continuous sampling process.

(iv) After each observation, without loss of time, the probabilities $p(a/x_1 \ldots x_n)$, $p(b/x_1 \ldots x_n)$ are calculated, where a and b are the two possible signals, $x_1 \ldots x_n$ the observed values of the signal presented.

(v) Decide for a if $p(a/x_1 \ldots x_n) \geqslant \dfrac{1-p}{1-p+p\mathrm{B}}$

Decide for b if $p(b/x_1 \ldots x_n) \geqslant \dfrac{p\mathrm{A}}{1-p+p\mathrm{A}}$, where

$p(b) = p$, $p(a) = 1 - p$, after any number, n, of observations.

(vi) The values of A and B are adjusted to produce the fastest average reaction time compatible with a given overall error level, ϵ.

$$\text{If } \sigma_a = \sigma_b = \sigma \text{ then } n = \frac{2\sigma^2}{(b-a)^2}\left[(1-2p)\log\frac{p}{1-p} + (1-2\epsilon)\log\frac{1-\epsilon}{\epsilon}\right],$$

which curve is slightly more bowed than that which information theory would predict. Both these curves are shown in Figure 2; the constants in the curve derived from the sequential probability model have been adjusted to produce coincidence with the information theory curve at $p = \frac{1}{2}$ and $p = \frac{1}{4}$, and hence to demonstrate the very close similarity of the two curves over that region of the frequency scale which has usually been studied.

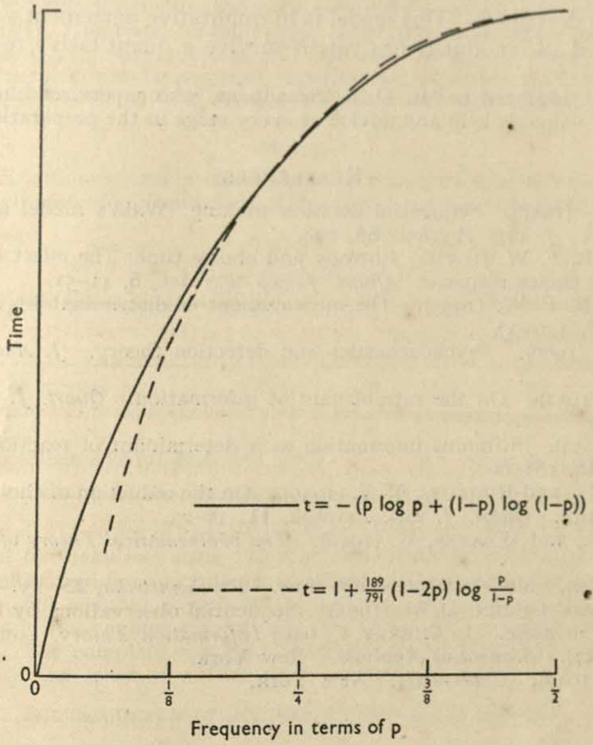
If we define $\mathrm{R}_\epsilon\left(\frac{b}{a}\right) = \dfrac{\text{no. of errors on presentation of } b}{\text{no. of errors on presentation of } a}$, then $\mathrm{R}_\epsilon\left(\frac{b}{a}\right) = \dfrac{1-p-\epsilon}{p-\epsilon}$,

and for $p = 0{\cdot}25$, and $\epsilon = 0{\cdot}05$, $\mathrm{R}_\epsilon\left(\frac{b}{a}\right) = 3{\cdot}5$, which is not so extreme an unbalance of errors as that observed in the experiments.

However, if we put $\sigma_b > \sigma_a$ the values of A and B to satisfy postulate (vi) are altered in such a way as to straighten the time-frequency curve and increase the value of $\mathrm{R}_\epsilon\left(\frac{b}{a}\right)$, provided $p < 0{\cdot}5$, as in Experiment II. Thus this model is in qualitative agreement with the results of the experiments described here. It will predict a longer reaction time to the infrequent signal, and a longer average reaction time when the majority of the signals are long as opposed to short. It also makes the prediction that if we repeated Experiment II with more long than short signals, we should find a time-frequency curve more bowed than that predicted by information theory, and a value of $\mathrm{R}_\epsilon\left(\frac{b}{a}\right) < \dfrac{1-p-\epsilon}{p-\epsilon}$.

One very important difference from earlier models (cf. the one discussed by Crossman (1955)) arises from the use of a sequential decision procedure. In a fixed sample model the decision process is terminated when a certain number of observations have been made. Consequently, unless some specific postulate to the contrary is made, reaction time will be constant, and in any case will be independent of the signal presented and the response made. In a sequential model, on the other hand, the decision process is terminated when a certain likelihood in favour of one or other of the alternatives has been reached. Thus the number of observations required will vary, and, in the case of the model described above, the response time will be approximately normally distributed by virtue of the decision process involved.

FIGURE 2



The time-frequency relation predicted by the model for two alternative signals distributed with equal variance, compared with the relation obtained from Information Theory.

As yet this model makes no prediction about the effects of practice; but a reasonable agreement with experimental results could probably be achieved by the addition of another postulate, such as supposing that the time required to make one observation, δt, decreases with practice. This model is capable of generalization to more than two alternatives, and explicit *n*-choice solutions have been derived. For the basic idea of the model the author is indebted in part to a paper by Stone (1960).

Wald's sequential probability ratio test has hitherto been applied as a possible decision model in situations involving a series of discrete presentations of the same signal by Swets and Green (1961), and Becker (1958) has used it in a different context altogether. Non-stochastic decision models have been applied to the problem of detecting sub-threshold visual and auditory signals; work in this field has been reviewed by Green (1960).

CONCLUSIONS

On the evidence presented in this paper, Information Theory, in the form of a constant average rate of transmission of information, without any coding delays, is untenable in the sort of situation studied here. It is true that $t \propto -\sum_i p_i \log p_i$ is frequently a good approximation to the time-frequency relation, but significant deviations from it have been obtained, and the time-frequency relation probably involves other factors, particularly error-rate and ratio.

A model, based on the sequential probability ratio test for the normal distribution, giving a hypothetical relationship between time, frequency and errors for constant ratio, has been described. This model is in qualitative agreement with most of the facts established so far, but it has yet to survive a quantitative test.

I am greatly indebted to Mr. D. E. Broadbent, who supervised these experiments, and gave much valuable help and advice at every stage in the preparation of this paper.

REFERENCES

BECKER, G. M. (1958). Sequential decision making: Wald's model and estimates of parameters. *J. exp. Psychol.*, **55**, 628.

CROSSMAN, E. R. F. W. (1953). Entropy and choice time: The effect of frequency unbalance on choice response. *Quart. J. exp. Psychol.*, **5**, 41–51.

CROSSMAN, E. R. F. W. (1955). The measurement of discriminability. *Quart. J. exp. Psychol.*, **7**, 176–95.

GREEN, D. M. (1960). Psychoacoustics and detection theory. *J. Acoust. Soc. Amer.*, **32**, 1189.

HICK, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

HYMAN, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

MOWBRAY, G. H., and RHOADES, M. V. (1959). On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

SHANNON, C. E., and WEAVER, W. (1949). *The Mathematical Theory of Communication.* Urbana.

STONE, M. (1960). Models for reaction time. *Psychometrika*, **25**, 251–60.

SWETS, J. A., and GREEN, D. M. (1961). Sequential observations by human observers of signals in noise. In CHERRY, C. (ed.) *Information Theory*, London.

WALD, A. (1947). *Sequential Analysis.* New York.

WIENER, N. (1948). *Cybernetics.* New York.

SHORTER ARTICLES AND NOTES

ON CHOOSING BETWEEN BETS

BY

H. C. A. DALE

From the Medical Research Council, Applied Psychology Research Unit, Cambridge

Most attempts to account for the way persons choose between bets emphasize the use which is made of information about their basic attributes, i.e. the gains and losses which might be incurred and their respective probabilities. In experimental investigations of betting behaviour this is usually the only information which the subjects are given (e.g. Edwards, 1955; Hurst and Siegel, 1956). Everyday observation suggests, however, that if given the opportunity subjects will also use information about the outcome of bets. In the experiments to be reported, experimental subjects were told what gains had accrued from three bets in the past, the purpose being to see how this information affected their betting.

METHOD

Thirty-six subjects were run in three groups of 12 (six experimental and six controls). All were repeatedly forced to choose one of the three bets shown in Table I. Play points served as currency. They sat at individual tables in four rows of three, the experimental subjects being in the back two rows. Subjects recorded their choices on betting slips, stamped to show their origin, which were collected on every trial. As each was collected the subject was told whether it had led to gain or loss. Eighty-five trials were run altogether. In addition, subjects were required at intervals to rank the three bets in order of worth. The complete sequence of events during the experimental session is shown in the following scheme:—

TABLE I

THE THREE BETS

	Worst	*Middle*	*Best*
Probability of success	0·2	0·5	0·8
Prize	7	5	3
Forfeit	3	5	7
Expectation [Mean overall gain per trial]	−1	0	+1

THE SEQUENCE OF EVENTS DURING EACH SESSION

Instructions.
Questionnaire I: in which subjects were required to rank the bets in order of worth.
Five Demonstration trials run without a fixed timing while the experimenter ensured that instructions were understood.
Information about the outcome of previous bets provided for experimental subjects.
Twenty trials at 45 sec. intervals.
Questionnaire II: ranking of the bets.
Twenty trials at 45 sec. intervals.
Memory test: Subjects examined on their memory of the outcome of the bets they selected.
Questionnaire III: as Questionnaires I and II.
Fresh start: all subjects told that the first forty trials were given for experience and that only those yet to come really counted.
Twenty trials at 45 sec. intervals.
Questionnaire IV: as Questionnaires I, II and III.
Twenty trials recorded *en bloc.*

The task was introduced as an exercise in decision taking. Care was taken to avoid all mention of gambling. The nature of risky decision taking was described by reference to the problem of timing the demolition of a bridge when retreating in battle. If it blows too soon friends are left behind, if too late the enemy can cross. It was stressed that good performance at decision taking was to be assessed only in the long run.

The outcome of each of the five demonstration trials was determined by drawing a counter at random from each of three cups containing black and white counters in proportions which corresponded to the probability of success for each bet. Thus on these trials the subjects learned the outcome of all three bets. On subsequent trials the prepared lists were used, though subjects were assured that the outcomes had been determined in the same manner. In actual fact, the lists had been prepared from random number tables in such a way that during the first and second sets of 20 trials the bets did not quite pay off according to expectation, whereas during the third set they concurred with it exactly. Two different lists were prepared. One was given to the first and third groups. The other to the second.

Information about the outcome of previous bets was shown to the experimental subjects by erecting a blackboard half-way down the laboratory so that it faced only the back two rows. It showed the gains which had accrued from each bet during a previous run of 40 trials. For the worst bet ($p = 0{\cdot}2$) it showed a loss of 40 points: for the middle bet ($p = 0{\cdot}5$), no change at all; and for the best bet ($p = 0{\cdot}8$), a gain of 50 points.

Most trials were run at 45 sec. intervals, but the last 20 were run *en bloc.* Each subject recorded all 20 on a sheet of paper, and then all were run off together. No knowledge of results was obtained, therefore, between these last trials.

RESULTS

From Table II it can be seen that whereas during the demonstration trials, there was no difference between the choices made by experimental and control subjects, on subsequent trials the experimental subjects chose the most profitable bet far more frequently than did the controls. Both groups showed evidence of learning, but the performance of the experimental subjects on their first ten trials was superior to that of the controls at the end of the session. The difference in the frequency with which they chose the best bet during the last 20 normal trials, i.e. blocks five and six, is highly significant ($t = 2{\cdot}7$ with 34 d.f.).

TABLE II

MEAN NO. CHOICES OF EACH BET ON SUCCESSIVE BLOCKS OF 10 TRIALS

Block	*The five demonstration trials* × 2	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Experimental subjects ($n = 18$)									
Best	4·5	6·4	7·4	8·4	8·1	8·4	8·4	8·6	8·6
Middle	4·8	3·2	2·1	1·6	1·4	1·3	1·3	1·1	1·2
Worst	0·7	0·4	0·5	0·0	0·5	0·3	0·3	0·3	0·2
Control subjects ($n = 18$)									
Best	4·7	5·2	4·8	5·2	5·3	5·7	5·6	6·2	6·2
Middle	4·3	3·7	3·7	4·1	3·8	3·7	3·7	3·5	3·3
Worst	1·0	1·1	1·6	0·8	0·9	0·6	0·7	0·3	0·5

Changes in the rating of the bets are shown in Table III and the relation between these ratings and performance in Table IV.

TABLE III

THE NUMBER OF SUBJECTS WHO CORRECTLY NAMED THE BEST BET

Questionnaire	*1*	*2*	*3*	*4*
Experimental subjects ($n = 18$)	9	17	16	17
Control subjects ($n = 18$)	12	13	12	13

TABLE IV

KNOWLEDGE AND PERFORMANCE

Questionnaire	*1*		*2*	
Best bet named:	*Correctly*	*Wrongly*	*Correctly*	*Wrongly*
Mean number of choices of best bet* . . made by experimental subjects (maximum = 10)	8·3	4·4	8·7	3·0
Mean number of choices of best bet* . . made by control subjects (maximum = 10)	6·4	2·8	7·4	1·6

* Performance on the first block of 10 trials is shown in relation to the ratings given on the first questionnaire: that on the fifth in relation to those on the third.

DISCUSSION

It is clear from Table II that the experimental subjects used the information about the outcome of previous bets to their advantage. Only one half of them had been able to rate the bets correctly on the basis of information about their basic attributes (see col. 1 of Table III) and it is likely that some of these had guessed and were therefore correct only through chance. Knowledge of the outcomes led them to alter their ratings and their choice behaviour, the two being closely linked. This strongly suggests that the initial response was made in ignorance and that, rather than attempt to account for a subject's initial response in a betting situation by asserting that he distorts the information provided and calculates subjective expectation, as many theorists attempt to do (Edwards, 1955; Hurst and Siegel, 1956), we should accept that he is really unable to make proper use of this information.

REFERENCES

EDWARDS, W. (1955). The prediction of decision among bets. *J. exp. Psychol.*, **51**, 201–14.

HURST, P. M., and SIEGEL, S. (1956). Prediction of decisions from a higher ordered metric scale of utility. *J. exp. Psychol.*, **52**, 138–44.

THE INFLUENCE OF THE NUMBER OF ALTERNATIVES ON THE PERCEPTUAL RECOGNITION THRESHOLD*

BY

P. FRAISSE and M. BLANCHETEAU

From the Laboratory of Experimental Psychology, Sorbonne, Paris

A study of the recognition thresholds for nonsense syllables previously presented in lists containing from 2 to 15 syllables is reported. The threshold is found to increase in proportion to the logarithm of the number of alternatives (i.e. the length of the list from which a given test syllable is drawn). It is therefore concluded that recognition threshold is a function of the amount of information transmitted.

It is easier to perceive an expected than an unexpected stimulus. In other words, an attitude congruent with an event sensitises the percipient to the event itself. The effects of the sensitisation are various: the most important is manifested in the *selectivity* of perception. According to our attitude, we perceive certain stimuli whereas others are missed or barely noticed. A preperceptive attitude can also affect the content of perception, yielding a compromise between what was given and what was anticipated. But among all these effects, there is one that is particularly noteworthy and which has been frequently stressed by psychophysicists. The absolute threshold is lower when one knows the nature of the stimulus which one is about to perceive. This fact, banal in itself, has however been little studied. In particular, we do not really know the law that governs the level of the visual perception threshold in relation to increasing dispersion of subjective attitude; i.e. when the number of alternatives which one might expect increases. True, Postman and Bruner (1949) showed that the minimum duration of stimulation necessary for a subject to recognize which word in a pair denoted a colour was 191 millisec. and that the threshold was raised to 225 millisec. if the subject's task was to recognize which word in each pair denoted *either* a colour *or* an article of food; that is to say, the threshold was raised by 37 millisec. when the subject had to anticipate a word belonging not to one but to two different verbal categories.

The studies closest to our own are those of Miller, Heise and Lichten (1951). These authors ascertained percentage of words recognized in relation to intensity of masking noise and number of alternatives given. The subject's task was to recognize a partially masked word in lists of 2, 4, 8, 16, 32 or 256 monosyllables. The curves representing percentages of words recognized in relation to the signal-to-noise intensity ratio expressed in decibels were sigmoid, and it was found that the amount of information necessary to recognize a given item is a function of the range of alternatives offered to the subject.

Bruner, Miller and Zimmerman (1954) have studied this same problem, but in relation to the degree of learning of lists of 8, 16, 32 or 64 monosyllabic words. They found that at all stages of learning the results were affected by the number of alternatives. After the first reading of the list, the recognition percentages were respectively 2 per cent., 1·75 per cent., 0·6 per cent. and 0·3 per cent., as the total was increased from 8 to 64 words. At the end of the training period, the percentages became

---

* The Editor wishes to thank Dr. Alice Heim and Dr. Roy Davis for their valuable assistance in translating this paper.

92 per cent., 74 per cent., 64 per cent. or 50 per cent. (These percentages have been calculated from the raw data given by the authors.)

It may be inferred from these studies that the greater the number of alternatives the higher the threshold.

We wished to take further the systematic study of this problem and in particular to investigate visual recognition thresholds to determine the law relating increase of threshold to number of alternatives given to the subject. Our hypothesis is that the perceptual recognition threshold is proportional to the amount of information transmitted or, put another way, that the difficulty of recognizing an item is proprotional to the amount of information contained in it.

EXPERIMENTAL METHOD

It was decided to use nonsense syllables in order to exclude interference due to the relative frequencies of words and their availability to a given subject. Thirty three-letter syllables (one vowel between two consonants) were chosen having a very low co-efficient of association (as determined by previous unpublished work from this laboratory). From this series, lists of 2, 4, 7, 10 and 15 syllables were made up. The lists of 2, 4, 7 and 10 syllables comprised different syllables; that of 15 syllables included certain syllables which appeared also in the shorter lists. From our selection of 30 syllables, six different series of lists were prepared such that each syllable was presented roughly an equal number of times.

For each list, one test syllable was chosen such as that over the whole six series of lists, each syllable was used once and once only. A single series of lists with the five levels of alternatives was used for each subject. The five lists of 2, 4, 7, 10 or 15 syllables were presented in a different order to each subject, such that the average ranking of each list was the same. Finally, as a last precaution, the test syllables were never in the first or last position, except of course with the list of two syllables.

The experiment was preceded by a preliminary practice run designed to familiarize the subject with tachistoscopic perception. This practice run was conducted as though part of the experiment proper: the subject was asked to read certain syllables, different from those used later, and told that he would have to recognize from among these one that would be shown to him in the tachistoscope, starting with a very brief exposure and increasing the exposure time from presentation to presentation until he recognized the item. A Dodge–Gerbrandt tachistoscope was used governed by an electronic timer. Each syllable, printed in indian ink on white bristol card, appeared in the centre of a rectangle 6·1 cm. wide by 5 cm. high. (The letters were 5 mm. high and the thickness of the lines was 0·5 mm.)

The experiment proper was then begun. In the first stage, the subject was required to familiarize himself with the 30 syllables. Each syllable was written on a card in the same characters as those used later for tachistoscopic presentation. The subject was asked to read the list five times, each time in a different order, with instructions to try to learn them by heart. The order of presentation of the syllables varied not only for each reading of the same subject but also as between the different subjects.

The second stage of the experiment was now begun: the subject was told that from the total collection certain syllables had been chosen, e.g. four if one begun with the list of four. From these four, he would be shown one and one only until he recognized it. Exposure time began at 6 millisec. and lengthened by intervals increasing in geometrical progression in the ratio 1:2. Between each tentative identification, i.e. between each presentation, the subject re-read aloud the list of chosen syllables (e.g. four in the example we have given). The subjects were 36 first-year psychology students. For each set of alternatives we calculated the mean and the standard deviation of the 36 thresholds thus determined.

RESULTS

The main results are given in Table I. In Figure 1, the mean threshold of recognition times shown in this Table are plotted against the logarithm of the number of alternatives. It seems legitimate, at first sight, to treat the regression line as a straight line, this interpretation being justified by the analysis given in Table II.

FIGURE 1

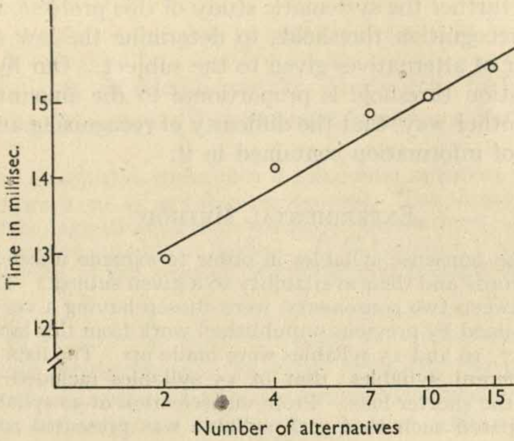


TABLE I

Number of alternatives	2	4	7	10	15
Mean threshold of recognition ..	12·9	14·1	14·8	15·0	15·4
Standard deviation	2·6	2·6	3·3	3·3	3·2

TABLE II

Source	*d.f.*	*MS*	*F*
Between subjects	35	47·03	13·1 $p < 0·01$
Between alternatives ..	4	33·59	9·35 $p < 0·01$
Linear regression* ..	1	126·80	35·3 $p < 0·01$
Other regression	3	2·53	0·7
Residual	140	3·59	—

* Calculated between thresholds and logarithm of number of alternatives.

It is thus plain that the threshold increased in proportion to the logarithm of the number of alternatives; that is, in accordance with Shannon's definition, as a function of the amount of information transmitted (Shannon, 1948).

Frick (1953) has established a comparable law, not for recognition thresholds, but for the time required to name a word. He was able to show a linear relation between increase in verbal reaction time and the logarithm of the number of words comprising the "ensemble size" from which the word to be recognized had been drawn (the numbers varied from 4 to 24). At the same time, these two results are not identical as one does not yet know the exact law relating the threshold of recognition to the verbal reaction time.

It is interesting to establish, however, that the influence of the number of alternatives appears to obey the same law in the case of both these types of perception. This proportion between the threshold of perception and the logarithm of the number of alternatives probably holds when there are a large number of alternatives. But one might ask, as has Bruner (1957, p. 136), if the same law would apply to the case where the number of alternatives is smaller than that of the span of apprehension and to the case where it is greater. In the case of 2, 4 and indeed 7 syllables, the subject might at the moment at which the stimulation occurs have all the alternatives simultaneously in consciousness, although even if the subject could have learnt the list by heart (which is practically impossible), one could not say that all the alternatives were simultaneously in consciousness. Such a difference in the availability of the alternatives could have strikingly decreased the threshold in cases in which the number of alternatives was fewer than 7. Our results render such an hypothesis improbable. The amount of information transmitted is the sole determining factor and upon this the availability of perceptual schemata directly depends.

REFERENCES

BRUNER, J. S. (1957). On perceptual readiness. *Psychol. Rev.*, **64**, 123–52.

BRUNER, J. S., MILLER, G. A., and ZIMMERMAN, C. (1955). Discriminative skill and discriminative matching in perceptual recognition. *J. exp. Psychol.*, **49**, 187–92.

FRICK, F. C. (1953). Some perceptual problems from the point of view of information theory. In *Current Trends in Information Theory*. Pittsburgh.

MILLER, G. A., HEISE, C. A., and LICHTEN, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J. exp. Psychol.*, **41**, 329–35.

POSTMAN, L., and BRUNER, J. S. (1949). Multiplicity of set as a determinant of perceptual behaviour. *J. exp. Psychol.*, **39**, 369–77.

SHANNON, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 379–423 and 623–56.

PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1961

12th–14th April, 1961. Extended meeting at Cambridge.

1st Session: "Illusions and constancy," by R. L. Gregory and C. D. Shopland. "Studies in concept identification," by E. J. Archer. "Binocular cues," by S. Papert. "The psychophysical law," by M. Treisman.

2nd Session: "Simultaneous brightness contrast," by Lord Burgh (by invitation). "Effects of age and experience on taste discrimination in the great tit," by M. Vince. Three short papers on stabilized retinal images by H. Barlow, G. Brindley and F. Campbell.

3rd Session: "Integration time and difficulty of discrimination," by I. Hughes (by invitation). "Individual differences in performance after loss of sleep," by D. Corcoran (by invitation).

4th Session: "The use of perimetry with monkeys in investigating the effects of brain lesions on visual behaviour," by A. Cowey (by invitation). "The physical measurement of human rhodopsin and its relation to dark adaptation," by W. A. H. Rushton (by invitation).

5th Session: "The two-colour threshold as a measure of normal versus defective colour vision," by R. M. Boynton. "Land colours, simultaneous contrast and colour theory," by C. E. M. Hansel (by invitation).

5th–6th July, 1961. Extended meeting at Oxford.

1st Session: "Games theory and reaction times," by J. Annett. "An experiment in phonetic symbolism in Tanganyika," by R. Davis.

2nd Session: "Binocular rivalry and stereoscopic depth perception," by A. Treisman. "An experimental investigation into nominal dysphasia," by M. Williams and G. Rochford (by invitation). "Attention; some theoretical aspects," by J. A. Deutsch.

3rd Session: "Two origins of achievement motivation," by J. M. Argyle (by invitation). "Drive studies with intraventricular injections," by J. Herberg (by invitation). "Timing eye and hand movements," by E. R. F. W. Crossman and G. R. Bunting.

4th Session: "Warning intensity and threshold change," by M. Treisman. "The verbal and non-verbal abilities of children with unilateral electroencephalgraphic foci," by M. Annett (by invitation).

3rd–5th January, 1962. *14th Annual General Meeting at Birkbeck College, London.*

1st Session: "The masking of a brief visual stimulus by a random pattern," by M. Kinsbourne and E. K. Warrington. "Visual 'noise' as a tool of perceptual research," by D. M. MacKay. "Guesses-to-recognize and guesses-to-recall," by J. Brown. "The influence of rate, vocabulary size and a supplementary task on the generation of random letter sequences," by A. Baddeley (by invitation).

2nd Session: "Factors affecting the duration of a short period spent in isolation," by R. C. B. Aitken and J. L. Gedye (by invitation).

3rd Session: "Stabilizing the retinal image," by C. Evans (by invitation). "Changes in arousal with stimulation," by P. Venables. "Stress, arousal and decision," by D. E. Broadbent.

4th Session: "Awareness of errors in immediate memory," by J. S. Wilfred (by invitation). "Drug-induced changes in motor skill and the perception of hand position," by D. Legge (by invitation). "Discrimination without awareness," by N. Dixon and A. Worthington.

5th Session: "Localization of a memory trace in one cortical hemisphere and transfer to the other hemisphere," by I. Russell (by invitation).

Committee, 1962: President: Professor H. Kay.

Editor: Professor O. L. Zangwill.

Ordinary Members of Committee: Dr. R. J. Audley (Hon. Secretary), Dr. J. Brown, Dr. R. Davis, Dr. R. Harper, Dr. C. Poulton, Dr. J. Symonds, Dr. J. Szafran (Hon. Treasurer), Dr. M. Treisman, Dr. P. Venables (Asst. Hon. Secretary).

The following have been elected into membership of the Society:

P. Bakan, Ph.D., Department of Psychology, Michigan State University, U.S.A.
E. Belbin, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
J. L. Gedye, M.B., B.Ch., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.
R. A. M. Gregson, Ph.D., 26 Thurlestone Avenue, London, N.12.
D. M. Holding, Ph.D., Department of Psychology, The University, Leeds, 2.
H. Ikeda, Ph.D., Department of Psychology, Maudsley Hospital, Denmark Hill, London, S.E.5.
I. Oswald, M.D., Department of Psychological Medicine, The University, Edinburgh.
W. A. H. Rushton, F.R.S., Physiological Laboratory, Downing Street, Cambridge.
W. Sluckin, Ph.D., Department of Psychology, The University, Leicester.
J. R. Smythies, M.D., Department of Psychological Medicine, The University, Edinburgh.
A. C. Staniland, Ph.D., Psychological Laboratory, Gandy Street, Exeter.
K. P. Watts, Psychological Laboratory, Downing Street, Cambridge.
M. Williams, D.Phil., Littlemore Hospital, Littlemore, Oxford.

Visiting Foreign Member:

Dr. S. Chorover, National Hospital, Queen Square, London, W.C.1.

BOOK REVIEWS

Annual Review of Psychology, Volume 12. Edited by P. R. Farnsworth, O. McNemar and Q. McNemar. Annual Reviews, Inc., Palo Alto, California. 1961. Pp. 533. $7.50.

Two years ago Sutherland found that the area in which steady progress had been most apparent was that of physiological psychology. This trend seems to have continued, although there is still no major breakthrough. (It is, indeed, difficult to see what could count as a major breakthrough in physiological psychology. Now that the nature of the nervous impulse is known, and the mapping of the functional anatomy of the nervous system is in full swing, about the only outstanding problem would seem to be the discovery of the nature of the changes in neurones which corresponds to learning.) The development which seems particularly noteworthy is the increase in our understanding of the functions of the different levels of the afferent pathways, and the distribution of the various aspects of the encoding of information as shared among them. It is interesting to recall that only 8 years ago the writer was still being taught that the cortex was the functional and anatomical "centre" of the sensory system, and the afferent pathways merely conveyed information to the cortex.

A welcome trend in neighbouring fields is attention to what we might call "cross-cultural" studies in comparative psychology: that is, studies are beginning to relate behaviour to morphology and natural history, and not regard comparative psychology merely as the descriptive comparison of the abilities of different species. If comparative psychology has any meaning, its task should surely be to relate the evolution of behaviour patterns to the evolution of the physical form of organisms, and thus should be much more closely tied in with genetics and the biological sciences other than physiology. For far too long the maintenance of the bridge has been left to the ethologists alone. The time would seem ripe for a considerable expansion of interdisciplinary work on the relation of behaviour to survival value and speciation. (It seems a pity, e.g. that an investigation into the effects of dietary deficiency during pregnancy in rats used no more relevant task for the offspring than to run mazes.)

Unnecessary jargon is still apparent in these fields: "auditory" may be a defensible adjective, but "audition," and above all "olfaction" and "gustation" are not merely clumsy but lacking in taste.

How much notice of comments is taken by those compiling subsequent reviews it is hard to say, but whether or not due to the attack delivered last year in this *Journal*, there are signs that the fog of jargon is beginning to lift, and it is becoming possible to understand easily quite a lot even in those fields about which the reader has no expert knowledge; which is as it should be in such a publication. Indeed, one article, that by Boilideau and Boilideau on motor skills, is a positive joy to read, and even includes a list of "nominations for raspberries in the annual rewards programme." Too often objectivity in scientific writing is taken to be equivalent to drabness of language: it is a pleasure to find an exception.

It is perhaps worth considering possible reasons why the fields of social and personality psychology should be less satisfactory. Even leaving out of account these developments in technique so unusual as to be alarming (there is one in particular which appears to involve two psychiatrists talking out *their* troubles while the patient listens!) there remains what seems to be a desperate desire to be "scientific," together with a certain naivete. For example, violations of the "sure-thing" principle in decision theory are not all that surprising when one recognizes that even success can become a bore if continued for a sufficient time and if the outcome is certain. As an example of misdirected scientific objectivity, Stevens' experiment on the value of money, claiming to have found a power law relation between its worth to the subject and its numerical value, seems fully to merit the comment of the reviewer that, "it would require a lot of evidence to convince many researchers that subjects, asked how much money is half as desirable as $10, would systematically answer something different from $5." This underlies the point made by another reviewer that while it is quite legitimate to look for lawful relations in social psychology, social perception should not be treated as a kind of psychophysics of social objects, but in terms of individual differences.

A further point which all too often goes unnoticed in studies in these fields is the dependence of social behaviour on cultural patterns. No doubt if any worker were taxed directly with this point he would admit its importance, but there is all too often in practice what amounts to a complete insularity of outlook. The result is to enunciate generalizations which are almost certainly unjustified. If the aim of psychologists is to produce theories applicable in the main only to their own culture, well and good. But a claim by a reviewer that certain work shows that as they grow up girls tend to become more masculine while boys become less feminine smacks of a highly culturally determined study, and some indication should be given in the review as to the extent of generalization which is intended. If, on the other hand, laws are being sought which are valid even for statements about different cultures, then cross-cultural data are absolutely essential. A start on these lines is indicated in the work of Mohanna and Argyle on the effect of communication patterns in groups in Cairo and Oxford, and the reviewer connects these with American studies.

This brings up a major source of worry, about the *Annual Reviews*, and also about the development of psychology as a whole. There are many practical difficulties about organizing cross-cultural studies, but there should be less difficulty about the cross-cultural transmission of information. But unless there is a fashionable trend (as, for example, Russian work at the present time), the references to literature in languages other than English is vanishingly small. In a total of 300 entries in the bibliographies of six sections of the *Annual Review*, picked at random, there were only two references to articles not written in English (excepting Russian). I just do not believe that all work published in foreign languages is *that* bad. The only alternative is that a lot of this work is not being read for the preparation of the *Annual Review*.

Now it is true that it is becoming increasingly difficult to keep up with the mass of research literature. A colleague recently remarked with some justification that unless something is done it will soon be quicker to do any experiment than to try to discover whether it has been done by someone else. But it is precisely because of this problem that the *Annual Review* is so important. It is upon its reviewers that we are forced to rely to a large extent to keep us informed of what other people are doing. At the moment while trying nobly it seems not to be succeeding. And it is vital that it should. I believe that the reviewers have to acquire the journals that they need for themselves, without help from the publishers. If this is true, it should be altered—even to the extent of having a translation service of some kind provided.

Further, what should be the strategy of research and publication? As long as a man's worth is measured by the physical weight of his research publications, the situation will worsen. The signs suggest that this size-value illusion is one that increases with prolonged experience of it. The summaries of several sections of the *Review* bear witness to the fact that people are worried. Irwin's concluding remarks are particularly good on the unnecessary invention of new techniques for their own sake (or, one suspects, for the sake of increased status for the inventor). Similarly, Taylor and Nevis remark of the proliferation of new tests in personnel selection that, "too many untested tests are too readily available to too many unqualified testers."

Several things would seem to be true. Firstly, too many workers are more interested in producing new gadgets or tests than in modifying those which have been tried and found to be fairly successful. Secondly, there is far too much publication of piecemeal research. It seems to the writer that nothing can save us from being lost in the rising flood of publications except a willingness to publish a collected series of results from time to time rather than a vast number of individual short papers. Lastly, there is the general problem of research strategy. As Shaw puts it in his review of Group Dynamics: ". . . it is distressing to note a tendency to ignore much of the literature relevant to the problem under consideration. . . . Authors all too often refer only to studies conducted in their own laboratories, and occasionally to those from one or two other laboratories. . . . Whatever the reasons, this is a disturbing trend which one can only hope will soon be reversed.".

This warning is one with respect to which action must be taken very soon if it is to be successful. The trouble is that the present incontinence of publication is tied up with professional respectability and academic status.

To end on a brighter note, it is very pleasant to see the appearance of two new chapters, on Behavioural Decision Theory and Aesthetics as the first fruits of the recent appraisal of policy by the editors. An occasional review on Cybernetics should be included from time to time, since the work in that field has interesting logical implications as well as relevance to theory construction and the analysis of the mechanisms underlying behaviour. In the meantime, the chapters on Behavioural Decision Theory and Aesthetics are both very welcome, even if one is a little inclined to comment on the latter in the words of Hodgson that,

"Reason has moons but moons not hers
Lie mirrored on her sea;
Confounding her astronomers
But oh! delighting me."

NEVILLE MORAY.

Handbook of Physiology. Section 1: Neurophysiology. Volumes 2 and 3. Editor-in-Chief: John Field. Section Editor: H. W. Magoun. Executive Editor: Victor E. Hall. American Physiological Society (Sole Agents: The Williams and Wilkins Co., Baltimore). 1960. Vol. 2: pp. 781–1438; Vol. 3: pp. 1440–1968 + Subject Index Vols. 1–3, pp. 1969–2013. Each volume $20.

The first volume of this noteworthy Handbook was reviewed in this *Journal*, 1960, **12**, p. 191. The second volume is devoted to motor mechanisms (17 chapters), introduced by Dr. D. Denny-Brown, and central regulatory activities (eight chapters), introduced by Professor F. Bremer. Among chapters of special interest to experimental psychologists are those on cerebral control of eye movements (Whitteridge), regulation of feeding and drinking (Brobeck) and the reticular formulation (French). Although the distribution of emphasis is a little uneven, this volume keeps up well the high standard of its predecessor. The third volume is concerned: first, with the neurophysiological basis of the higher functions of the nervous system (roughly, physiological psychology) and contains 10 chapters, with a short introduction by Dr. W. Penfield; secondly, with the cerebral circulation and intracranial fluids (two chapters, with introduction by Dr. C. F. Schmidt); and thirdly, with neural metabolism and function (seven chapters, with introduction by Sir Rudolph Peters). Dr. R. W. Gerard contributes a final chapter on Neurophysiology: An Integration, in which he makes interesting reference to the possibility of translating into neurophysiological terms models derived from purely psychological studies and expressed in the language of information flow.

The large section on physiological psychology contains chapters on sensory discrimination (Neff), learning (Galambos and Morgan), drive and motivation (Stellar), emotional behaviour (Brady), perception (Teuber), attention, consciousness and sleep (Lindsley), thinking and memory (Halstead), skilled movement (Paillard) and speech (Zangwill). It thus provides an up to date review of the more important modern work bearing on these various topics—and of some of the older work too. The emphasis varies from mainly neurophysiological (Lindsley) to mainly psychological (Teuber). To the present reviewer, the inclusion of topics such as these in a book such as this is welcome evidence that the analysis of behaviour is now being taken seriously as a method in neurophysiology—indeed, we may even be entering the era of psychological physiology! At the same time, certain of the psychological chapters might have gained from an attempt to cover less ground and to focus on the contribution of psychology to physiology rather than *vice versa.*

As in the first volume, one must again deplore the paucity of reference to studies emanating from the Soviet Union. Although Galambos and Morgan deal seriously and at some length with Russian work, a chapter on conditioned reflexes by one of the more responsible Russian neurophysiologists would have filled an obvious gap. None the less, Dr. Magoun and his contributors will long stand in the debt of all serious students of the nervous system and merit warm congratulation on the high standard of their enterprise.

O. L. ZANGWILL.

An Introduction to Experimental Design. By W. S. Ray. The Macmillan Company, New York and London. 1960. Pp. x + 254. 45s. 6d.

This book gives an account of the simpler forms of experimental design which could be understood by any experimenter, however little mathematical knowledge he might have; a series of imaginary experiments, with simple numerical results, is used to demonstrate the uses of replication, analysis of variance and covariance, matching and factorial designs. The assumptions made in using significance tests for the variance cases are stated and the author makes sensible remarks about the need for randomization and the ways in which fallacious matching may occur. It would have been helpful, however, if he had indicated the amount of previous statistical knowledge that he required of his readers. A somewhat longwinded explanation of the use of sigma for summation is given in Chapter 4, yet Chapter 5, on random variation, would be incomprehensible to anyone not already familiar with the usual statistical methods and these could hardly avoid knowing about summation notation; the F-ratio is awarded some cursory remarks, the chi-square distribution is introduced into later arguments without any explanation of what it is. There is no explicit definition of either an observed or a theoretical distribution. Parameters (see page 41) are usually defined to be constants which occur in the specification of a theoretical distribution; the definition given here implies that there is an infinite number of them for each distribution. Placing a larger variance over a smaller will inevitably give a ratio greater than 1; the reason given for halving the probability values (page 58) could thus hardly be understood by anyone meeting it for the first time.

Throughout the book the statistical discussion is at a much more sophisticated level than is the explanation of the experimental designs. A newcomer to statistics would need to make use of some simpler book on elementary statistical methods. Nevertheless, he should find this introduction extremely useful in some ways. Although the book is aimed at psychologists, there is nothing specifically psychological about the contents, apart from the terms in which the experiments are described, and students in other fields may find it valuable, with the same reservations. V. R. CANE.

The Informed Heart: The Human Condition in Modern Mass Society. By Bruno Bettelheim. London. Thames and Hudson, 1961. Pp. viii + 309. 35s. net.

The author of this book was trained as a psychoanalyst in Vienna and spent a year in Dachau and Buchenwald shortly before the last war. His experiences led him to write a remarkable article on human behaviour in the concentration camps which was published in the *Journal of Abnormal and Social Psychology* in 1943; this article, considerably expanded, forms the core of the present book. Dr. Bettelheim has also some interesting things to say on the limitations of psychoanalysis as a social philosophy. Although not very well written, this is a frighteningly honest book and should be widely read.

O. L. ZANGWILL.

Van Nostrand Insight Books. No. 2. *Instinct*. Pp. ix + 181. No. 3. *Reinforcement*. Pp. viii + 230. No. 4. *Color Vision*. Pp. viii + 214. Edited by Robert C. Birney and Richard C. Teevan. London. D. Van Nostrand Co. Ltd. 1961. Each 11s. 6d.

This is a new series of American psychological paperbacks under the general editorship of D. C. McClelland. The three volumes under review, which have been edited by R. C. Birney and R. C. Teevan, are each intended to present a particular problem of enduring interest in terms of selected original articles arranged in historical sequence. As everyone who has tried to do something of the kind knows, this is an extraordinarily difficult thing to do, calling as it does for both scientific and historical judgement of a high order. Obviously, too, the success achieved will depend largely on the topic and on the degree of continuity manifest in its history. So far as the present trio go, colour vision is evidently in the strongest position and here the editors have produced an admirable volume. From Young through Helmholtz and Hering to Hecht, Granit, Willmer and Land, there has been real continuity in subject-matter, experimental method and level of discussion in spite of passionate differences in theoretical conviction. Instinct, on the other hand, scarcely lends itself to such historical compression. The idea has multiple roots in mental philosophy as well as in biology and multiple manifestations in psychological theory. Here, the line followed is roughly that from James through the Behaviourists to the Ethologists (though surprisingly there is nothing from Thorndike) with exclusion of other lines which run parallel, such as that from Darwin to McDougall and Freud. Although the present selection is not at all bad, it can certainly not be regarded as an adequate historical presentation of the instinct problem.

And what are we to say of Reinforcement? This is an All-American Special devoted to post-Thorndike behaviour theory; and it is a little odd, incidently, to find a work bearing this title described as an "Insight Book." There are selections from Hull, Mowrer and Tolman, and extracts from a number of recent experimental reports which bear on "drive reduction" and kindred issues, among them Olds and Milner's paper on intracranial "self-stimulation". Several of Dr. Neal Miller's papers are reprinted.

Although the standard of these volumes is quite high, it is a little difficult to see why it should be thought so necessary to serve students with old wine in new bottles (or should one say new cartons?). Most of the papers and extracts reprinted in this series are readily accessible, provided only that the student has the initiative to withdraw a periodical volume (other than the current issue) from the library shelf. Is this too much to ask of a University education?

O. L. ZANGWILL.

Pain: by F. J. J. Buytendijk. London. Hutchinson. 1961. Pp. 190. 25s.

To the experimental psychologist or physiologist, much of this book is likely to be irrelevant. It was first published in Dutch in 1943, and the considerable amount of scientific and medical work on pain that has been published in the following years is covered in an Annexe of eight pages. It has no claims, therefore, as a work of reference and in this brief attempt at up-to-dateness it is necessarily selective, as in paying only scant attention to the work of Hardy, Wolff, and Goodell, and in ignoring the admittedly controversial studies of Weddell and his co-workers.

The specialist will already be familiar with Beecher's very comprehensive review on "The Measurement of Pain" (1957), which covers 687 references, not 687 pages as stated in the Annexe. Both the specialist and the nonspecialist experimenter, or clinical worker, may find that the labour of following the philosophical viewpoints of such writers as L. Klages is not rewarded by the psychological outcome, namely that ". . . pain appears always in a certain *context* that determines the significance of the subjective experience."

K. R. L. HALL.

Selected Readings on the Learning Process. Edited by Theodore L. Harris and Wilson E. Schwahn. London. Oxford University Press. 1961. Pp. xi + 428. 28s.

This is a collection of papers on human learning, designed to supplement textbooks of Educational Psychology. The Editors argue that reading reports of actual experiments will help to give meaning to "the typical generalized text approach." This is surely true and such a book should therefore serve a wholly admirable function.

There are three parts. In the first there are a few attempts to define learning and to describe it in general terms; this is mercifully short. Next, there are a number of papers dealing with particular kinds of learning; there is a chapter called "Conceptual Learning," which includes Hull's "The Evolution of Concepts"; and another on "Skill Learning," which includes Book's experiments on typewriting. The third part deals with the influence of various factors upon learning; motivation, maturation, presentation of material, etc.; one chapter includes a short article by Piaget, which makes an admirable introduction to his work.

The actual selection will meet its purpose if it appropriately parallels the textbooks; this it no doubt does, since most of the articles are well known. But editors have other responsibilities, and one of them, in a book of limited size like this, is abridging; this they do rather crudely. Everything is abridged; even the short article by Piaget is made shorter. And often they omit parts so arbitrarily that it is difficult to make sense of what is left. For instance, in Hull's original monograph he tells us what the main measure of concept formation is to be and then continues "In addition there were three supplementary measures. . . ." In this book the main measure appears two pages *after* the supplementary measures, and "in addition" refers back to a description of the apparatus. This is seriously incompetent, but worse, although we are told that the experiment is to investigate the efficiency of proceeding from simple to complex as against complex to simple, the vital section explaining the nature of the complexity is left out. The one redeeming feature is that, in their confusion, students will be forced to refer to the original.

The Editors assume the further responsibility of writing introductions to the various sections. They begin by explaining that the book is not just about learning, but about the learning process, which is to be distinguished from the learning product (i.e. what is learned). But in the present context "learning," unlike "painting," is not ambiguous; presumably the opportunity for drawing a scientific distinction is not to be missed. Later, they go on about the supposed benefit to teachers of understanding "the pattern and sequence of change in behaviour and the functional concepts which illuminate key aspects of behavioural change." No doubt whatever is meant is important; but it is also important to realize (which the Editors don't appear to) that most of the experiments in this book will be of little help to an intelligent teacher. For the principles that are suggested, are already suggested by common sense, and their practicability can only be determined in the classroom.

The prose condemns itself: "Indeed, if the work of the teacher is to be effective, the importance of a high degree of professional competence in evaluating those aspects of the learning process which determine success or failure in learning cannot be underestimated." Communication is at such a low ebb that it hardly seems to matter that this is a contradiction and the opposite of what was intended.

But all this can be ignored; most of the articles are important, and, though carelessly abridged, it is useful to have them so readily available.

A. W. STILL.

PUBLICATIONS RECEIVED

NEW JOURNALS

Medicine, Science and the Law. Vol 1, Pt. 2, 1961. Edited by Francis E. Camps. Official Journal of the British Academy of Forensic Sciences. Published by Sweet & Maxwell, Ltd., New Fetter Lane, London, E.C.4. Annual Subscription £3 3s. Each part 17s. 6d.

Darshana: An International Quarterly of Philosophy, Psychology, Psychical Research, Religion, Mysticism and Sociology. Editor: J. P. Atreya. Published in Moradabad (India). Annual Subscription £1.

Journal of Psychiatric Research. Vol. 1, No. 1, Oct., 1961. Editor-in-Chief: Seymour S. Kety, Asst. Editors: Joel Elkes and David A. Hamburg. Published by Pergamon Press. Annual Subscription (A) for libraries: £7, (B) for private subscribers: £3 10s.

Problems of Psychology. Nos. 1 and 2. 1960. Hon. Scientific Translation Editor N. O'Connor. Published quarterly by Pergammon Press, containing selected papers translated from Voprosy psikhologii. Annual subscription: (A) for libraries £15, (B) for private subscribers £4.

BOOKS AND OTHER PUBLICATIONS

Review in a later issue is not precluded by notice here

The Psychology of Expression. By Sylvia Honkavaara. British Journal of Psychology Monograph Supplement No. 32. London. Cambridge University Press. 1961. Pp. x + 96. 25s.

Mental Health Research Institute Fourth Annual Report. July, 1959–*June,* 1960. University of Michigan. 1960. Pp. 62.

Psychodiagnostisches Vademecum: Hilfstabellen für den Rorschach-Praktiker. By Ewald Hohm. Bern. Hans Huber. 1960. Pp. xv + 168. Fr./DM 22.80.

General Systems: Yearbook of the Society for General Systems Research Vol. V. Edited by Ludwig von Bertalanffy and Anatol Rapaport. Ann Arbor. Society for General Systems Research, 1960. Pp. xix + 235. $7.50.

Rorschach Psychology. Edited by Maria A. Rickers-Ovsiankina. London. John Wiley. 1960. Pp. xvi + 483. 68s.

Letters to My Teacher. By Dagobert D. Runes. New York. Philosophical Library Inc. 1961. Pp. 105.

Attitude Organization and Change: An Analysis of Consistency Among Attitude Components. By Milton J. Rosenberg, Carl I. Hovland, William J. McGuire, Robert P. Abelson and Jack W. Brehm. Yale Studies in Attitude and Communication, Vol. 3. Yale University Press. 1960. Pp. xi + 239. $6.00.

Stuttering and What You Can Do About It. By Wendell Johnson. London. Oxford University Press (for Minnesota University Press). 1961. Pp. ix + 208. 32s.

Le Demon et Son Image. By Erwin Reisner. (Translated from the German by Jean Viret.) Bruges. Desclée de Brouwer. Textes et Études Anthropologiques. 1961. Pp. 237. 135 FB.

Studies in Mathematical Psychology I: Probabilistic Models for Some Intelligence and Attainment Tests. By G. Rasch. Denmark's paedogogiske Institute. Copenhagen. 1960. Pp. xiii + 184.

Psychology: An Introduction to the Study of Human Behaviour. By Henry Clay. Lindgren and Donn Byrne. London. John Wiley. 1961. Pp. xi + 429. 52s.

Industrial Productivity and Motivation. By H. C. Ganguli. London. Asia Publishing House. 1961. Pp. xiv + 92. 12s. 6d.

Handicapped Children. By John D. Kershaw. London. Heinemann. 1961. Pp. x + 228. 21s.

Psychometric Analysis of the Thematic Apperception Test. By R. Olavi Vittamäki. Finnish Academy of Science and Letters. Helsinki. 1959. Pp. 61.

Guildhall Lectures 1960. By Professor George W. Beadle, Professor H. J. Eysenck and Lord Adrian. The British Association/Granada. London. University of London Press. 1961. Pp. 105. 4s. 6d.

La vision en lumiere intermittente. By H. Pieron. Monographies Francaisse de Psychologie No. 8. Centre National de la Recherche Scientifique. Paris. 1961. Pp. 91. 8 NF. 50.

Persuasion and Healing: A Comparative Study of Psychotherapy. By Jerome D. Frank. London. Oxford University Press (for Johns Hopkins Press). 1961. Pp. xiv + 282. 35s.

Markov Learning Models for Multiperson Interactions. By Patrick Suppes and Richard C. Atkinson. London. Oxford University Press (for Stanford University Press). 1960. Pp. xii + 296. 66s.

Praktisches Handbuch Des Rorschach-Tests. By M. Loosli-Usteri. Bern. Hans Huber. 1961. Pp. 187. Fr./DM 22.50.

Motivation and Emotion: A Survey of the Determinants of Human and Animal Activity. By Paul Thomas Young. London. John Wiley. 1961. Pp. xxiv + 648. 86s.

Measurement and Evaluation in Psychology and Education (Second Edition). By R. L. Thorndike and E. Hagan. London. John Wiley. 1961. Pp. viii + 602. 86s.

Contemporary Issues in Thematic Apperceptive Methods. By J. Kagan and G. S. Lesser. Oxford. Blackwell Scientific Publications Ltd. (Springfield, Illinois. Charles C. Thomas.) 1961. Pp. xiv + 328. 96s.

A Ciba Foundation Symposium on The Nature of Sleep. Edited by G. E. W. Wolstenholme and Maeve O'Connor. London. J. & A. Churchill Ltd. 1961. Pp. xii + 416. 50s.

Contemporary European Psychiatry. Edited by Leopold Bellak. London. Evergreen Books Ltd. 1961. Pp. xxvi + 372. 25s.

Transactional Analysis In Psychotherapy. By Eric Berne. London. Evergreen Books Ltd. 1961. Pp. 270. 25s.

Group Psychotherapy with Children: The Theory and Practice of Play-Therapy. By H. G. Ginnott. London. McGraw-Hill. 1961. Pp. xvi + 208. 46s.

Determinants of Infant Behaviour. Edited by B. M. Foss. London. Methuens. 1961. Pp. xv + 308. 42s. net.

Behaviour. By D. E. Broadbent. London. Eyre & Spottiswoode. 1961. Pp. 216. 21s.

Sensory Communication. Edited by Walter A. Rosenblith. Published jointly by M.I.T. Press, Mass., and Wiley & Sons, New York and London. 1961. Pp. xiv + 844. 128s.

The Search for Meaning: A New Approach in Psychotherapy and Pastoral Psychology. By A. J. Ungersma. London. George Allen & Unwin Ltd. 1961. Pp. 188. 32s.

Verbal Learning and Verbal Behavior. Edited by Charles N. Cofer. Mc-Graw-Hill Series in Psychology. London. 1961. Pp. vi + 241. 46s.

Conditioning and Learning. By E. R. Hilgard and D. G. Marquis. (Revised by G. A. Kimble.) Methuen. London. 1961. Pp. ix + 590. 602.

Arbeitsanforderung und Berufseignung. By Von Heinz Schmidtke and Hugo Schmale Verlag Hans Huber. Bern & Stuttgart. 1961. Pp. 87. Fr./DM 13.50.

The Psychology of Persecution and Prejudice. By Dr. David Stafford-Clark. The Council of Christians and Jews. London. 1961. Pp. 27. 2s.

The Marquis de Sade. A Biography. By Gilbert Lély. (Translated by Alec Brown.) Elek Book Ltd. London. 1961. Pp. viii + 464. 50s.

The Historical Development of British Psychiatry. Volume I. 18th and 19th Centuries. By Denis Leigh. Pergamon Press. Oxford. 1961. Pp. xiv + 277. 70s.

Conceptions of Perceptual Defence. By W. P. Brown. British Journal of Psychology Monograph Supplement No. 35. Cambridge University Press. London. 1961. Pp. viii + 106. 25s.

Encephalitides. Proceedings of a Symposium on the Neuropathology, Electroencephalography and Biochemistry of Encephalitides, Antwerp, 1959. Edited by L. Van Bogaert, J. Radermecker, J. Hozay and A. Lowenthal. Elsevier Publishing Company. London. 1961. Pp. xiii + 721. £8.

Psychology of Judgement and Choice. By F. Restle. John Wiley. London. 1961.

A Model of the Mind. By Gerald S. Blum. Wiley. London. 1961. 53s.

Religious Experience and Truth. A Symposium Edited by Sidney Hook. New York University Press. New York. 1961. Pp. xiii + 333. $5.00.

Speech Disorders. By Sir Russell Brain. Butterworths. London. 1961. Pp. v + 184. 42s.

Hamdard Medical Digest. Vol. 5, No. 2, Feb., 1961. Edited by Hakim Mohammed Said. Hamdard (WAQF) Laboratories, Karachi, Pakistan.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIV 1962 Part 2

A TEST OF TWO THEORIES OF SHAPE DISCRIMINATION

BY

P. C. DODWELL

From Queen's University, Kingston, Ontario

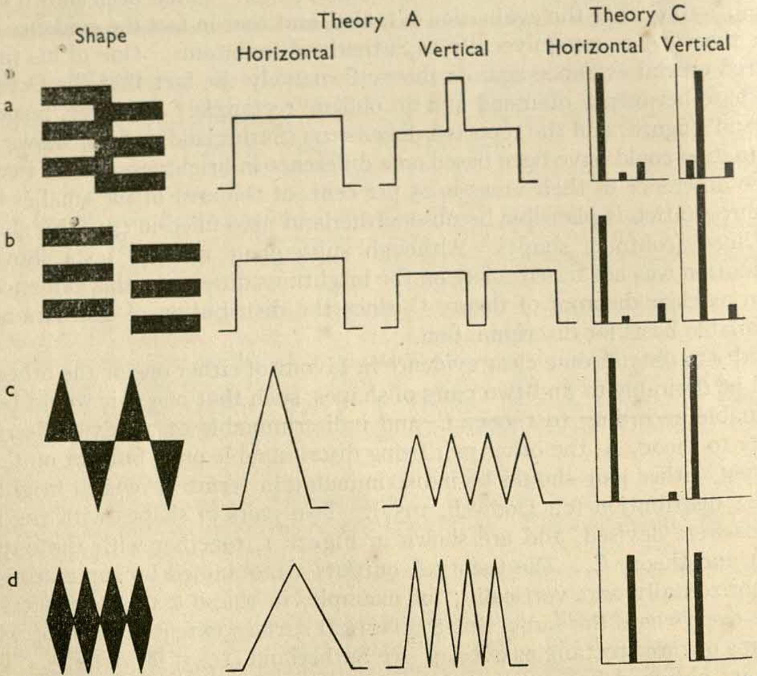
An experiment is reported which confirms a prediction from the writer's theory of shape recognition. The same experiment is held to refute an alternative theory, but this refutation is not conclusive, since the alternative theory has not been clearly stated and can be interpreted in several conflicting ways. It is shown that these possible interpretations of the alternative theory all yield predictions which are not confirmed by the present findings.

INTRODUCTION

In a recent evaluation of theories of shape discrimination, Sutherland (1960*b*) examined his own quite extensive findings on *Octopus* in relation to three theories. The theories are: A, his own theory (Sutherland, 1957); B, a recent theory of Deutsch's (1960); and C, a theory of the writer's (Dodwell, 1957). It has been shown elsewhere (Dodwell, 1961 *b*) that the evaluation is faulty, and that in fact the evidence does not support theory A as unequivocally as Sutherland maintains. One of his findings is considered crucial evidence against theory C, namely the fact that the Octopus can discriminate between a diamond and an oblique rectangle. However, inspection of Sutherland's figures and the reported dimensions (Sutherland, 1960 *a*) shows that the discrimination could have been based on a difference in brightness for the two shapes, since the difference in their areas is 25 per cent. of the area of the smaller of them. This interpretation is plausible because Sutherland used filled-in (surface) shapes and not outlined (contour) shapes. Although subsequent transfer tests showed that discrimination was not based solely on the brightness difference, this evidence cannot be taken as clear disproof of theory C, since the distribution of contours is not the only available basis for discrimination.

In order to obtain some clear evidence in favour of either one or the other theory, it would be desirable to find two pairs of shapes, such that one pair would be readily discriminable according to theory C, and indiscriminable or nearly indiscriminable according to theory A, the other pair being discriminable on A but not on C. To be a good test, either pair should be indiscriminable in terms of overall brightness, or brightness distribution (cf. Dodwell, 1957). Two pairs of shapes with the required properties were devised, and are shown in Figure 1, together with the outputs for theory A and theory C. The theory A outputs are obtained by summating surface extent horizontally and vertically; for example, in shape *a* the horizontal surface extent is everywhere the same, but the vertical surface extent is doubled where the two groups of three rectangles overlap (see Sutherland (1957) for details). Theory C outputs are obtained by counting the extent of contours in horizontal and vertica, directions: the outputs show, on the ordinate amplitude of discharge, on the abscissal the time at which the discharge occurs. For a discussion of how the amplitude and latency are calculated, see Dodwell (1957); they depend on the number of contours intersecting a horizontal or vertical line drawn at any given position on a shape. For example, in the horizontal output for shape *c* (theory C) there is a large discharge with small latency, followed by a small discharge with large latency. This is so because, over all of the shape except the horizontal "centre line," there are just four contours on any horizontal line (a ruler placed horizontally over the shape will intersect four contours), but at the centre line there is a single horizontal contour extending across the whole shape. On the other hand the horizontal output for *d* consists of a single large discharge, with twice the latency of the corresponding discharge for *c*, since any and every horizontal line drawn across the shape will intersect eight contours. Theory A predicts that the pair *a*, *b*, should be readily discriminable, since the vertical outputs differ. Theory C, on the other hand, predicts that the pair should only be discriminable with difficulty, since the outputs for the pair are quite similar, except for the additional discharge in the vertical output for *a*. Theory C predicts that *c* and *d* should be easier to discriminate since, apart from an additional discharge in the horizontal output for *c*, the amplitudes of the corresponding discharges differ, as do their latencies. For instance the latency for *d* (horizontal) is twice that for *c* (horizontal). Whilst the latencies for the principal vertical discharges are the same (since a vertical line would intersect just two contours at each point on each figure) their amplitudes differ. Theory A predicts that the pair *c*, *d*, should be more difficult to discriminate than the pair *a*, *b*, because of the overall similarity of the outputs for *c*, *d* (see, however, Discussion below).

FIGURE 1



The shapes used in the experiment, and the outputs for these shapes in theory A (Sutherland's) and theory C (Dodwell's).

EXPERIMENT

Aim

To assess the relative difficulty for the rat of discriminating between the pairs of shapes shown in Figure 1.

Apparatus

A modified form of Lashley jumping stand was used, as previously described (Dodwell, 1957). The shapes were cut out of white paper and pasted on to matt black cards. The total horizontal extent of any shape was not greater than 2¼ in., the vertical extent not greater than 2 in.

Subjects

The subjects were eight male hooded rats of the Long-Evans strain, all approximately 6 months old at the start of the experiment. They were maintained on a water deprivation schedule, having access to water for 4 hr. each day, and being run after approximately 20 hr. deprivation. The schedule had been in operation for about 3 months before the experiment began. During that time the rats had been taught various discriminations; first, a brightness discrimination without orientation or shape cues (Dodwell, 1960), then discrimination between horizontal and vertical striations, followed by discrimination of various simple geometric shapes.

Method

The subjects were divided into two groups of four, matched on the basis of their performance on the original brightness discrimination. Group I was trained on discrimination of the pair of shapes, *a*, *b*, in Figure 1, group II was trained on the pair *c*, *d*. No rats which had previously been trained with horizontal striations positive were put in group II. The reason for this was that *c* contains a horizontal contour as one of its "distinctive" features for theory C, so that responding to *c* could be a matter of straightforward transfer for rats previously trained with horizontal striations positive (Dodwell, 1957). Group I contained rats which had previously been trained on either horizontal or vertical positive, and there was no difference between the performance of the subjects with the two different antecedent conditions. Each subject was run for 25 days, 10 trials a day, and the numbers of correct and incorrect jumps recorded.

Results

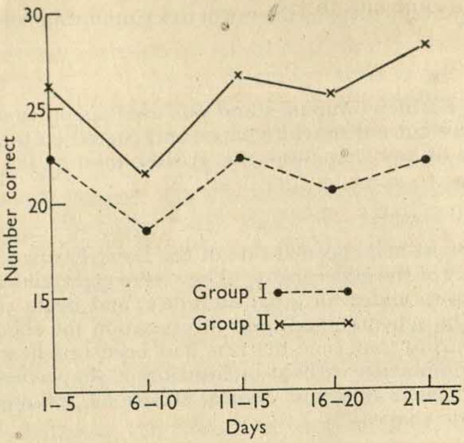
Performance in group II was superior to performance in group I. The overall course of learning is shown in Figure 2, where scores have been summed over periods of 5 days, so that each point is the average of 20 scores. Table I gives a trend analysis based on the scores (number of correct responses) for each rat for each day, following the procedure described by Edwards (1960). The F ratio for conditions shows that the overall superiority of group II is significant. The F ratio for trials (overall increase in correct responses with trials) just fails to reach significance at the 5 per cent. point, for which a ratio of 1·59 would be required. The conditions × trials interaction is not significant, indicating that there is no difference in trend for the two groups.

TABLE I

ANALYSIS OF LEARNING SCORES

Sources	*SS*	*DF*	*MS*	*VR*
(Subjects) . .	(68·3)	(7)		
Conditions . .	50	1	50	16·1 ($p < 0{\cdot}01$)
Residual (*a*)	18·3	6	3·1	
Trials	86	24	3·6	1·57
(S × T) . .	(366·4)	(168)		
C × T . .	36·7	24	1·5	
Residual (*b*)	329·7	144	2·3	
Total	520·7	199		

FIGURE 2



Proportions of correct jumps for rats trained on the shapes *a* and *b* of Figure 1 (group I) and those trained on shapes *c* and *d* of Figure 1 (group II). Each point represents the average of four rats for 5 days: the chance expectancy is 20 correct.

DISCUSSION

The results support theory C, since it was predicted that group II should discriminate better than group I. The analysis of variance indicates that there is not a significant increase in the number of correct choices with increased training, although the F ratio approaches the value required at the 5 per cent. confidence point. This is no doubt due to the fact that both groups showed some initial preference for their positive shapes, and this tendency decreased over the first few days (Fig. 2). The fact that there was some initial preference for the positive shapes suggests that the rats may have been transferring from previously learned discriminations. It is unlikely that this was so, for two reasons: (i) it has been found that there is generally a tendency for transfer effects to increase over the first 50 trials or so (Dodwell, 1960, 1961 *a*) and the initial preferences observed in the present experiment decreased; (ii) on either theory it would be expected that group II (all pretrained with horizontal striations negative) might have some tendency to treat *c* as a negative shape, although the reasons for this are quite different in the two theories.

Inspection of the outputs in Figure 1 seems to show that the outputs for *c*, *d*, differ quite as much as they do for *a*, *b*, on theory A, so that it might be argued that the prediction made from theory A is incorrect. This argument could only be sustained were it possible to state exactly how the outputs for theory A are analysed, compared or "recognized," but it is not possible to do this. Two suggestions have been made about the analysis of outputs for theory A: (i) that the analyser would ". . . analyse and compare the size of outputs from points on rows and columns" (Sutherland, 1957): (ii) that the analysis ". . . may take the form of an analysis of the rates of change of firing in the neurons whose firing pattern represents the horizontal and vertical projections of the original shapes" (Sutherland, 1958 *a*). It is clear from the way outputs are computed in theory A that the "sizes" (total amounts) of the horizontal and vertical outputs for any shape are always the same, since either output is an analogue of the area of the associated shape. Sutherland evidently intends the analysis to be based on the maximum height of the outputs, since he states elsewhere: "It is proposed that Octopus discriminates orientation by means of a neurological mechanism which counts the lateral and vertical extents of a shape, takes the ratio of one to the other, and classifies shapes by means of these ratios" (Sutherland, 1958 *b*).

A further possibility suggested by Sutherland is that a subsidiary system might compute the ratio of total contour to area for each shape. For such a system the pairs of shapes in Figure 1 should both be discriminable without difficulty.

The initial prediction must be modified in the light of these alternatives. There are actually four cases to consider for theory A.

(1) Discrimination is based on the ratio of the maxima for horizontal and vertical outputs for each shape, in which case *a*, *b*, should be about as discriminable as *c*, *d*. For the rat they are not and it seems unlikely that this ratio could be the basis for discrimination, since it would imply that all figures with equal vertical and horizontal extents, such as square, circle, diamond, any triangle in which the base and height are equal, etc., would be treated as equivalent; this is not the case for the rat or the octupus.

(2) Discrimination is based on rates of change and changes in rates of change in the outputs. This type of analysis would yield the same prediction as theory C *if* the analysis took no account of directions of change at points of discontinuity. It is not possible, however, that Sutherland intends this restriction to apply, since if it did some of the evidence he has claimed to be in favour of theory A would actually tend to refute the theory (e.g. Sutherland (1960 *b*), Figure 1, discriminability of shapes *f* to *m*).

(3) Discrimination is based on the ratio of contour to area for each shape. On this interpretation the pairs of shapes in Figure 1 should again be about equally discriminable, and since they are not, it seems that this ratio is not a basis for discrimination in the rat.

(4) Discrimination is based on the "shape" of the outputs, or the occurrence of peaks and troughs in the outputs in particular positions. This seems to be the intended basis for analysis, and it is an imprecise one. Despite this lack of precision, it seems clear that the initial prediction from theory A should hold, and the pair *a*, *b*, should be more discriminable than *c*, *d*, because there is a peak in the vertical output for *a* where a trough occurs for *b*, but no such marked difference for the other pair.

Of the four possibilities, (1) and (3) predict that the two pairs of shapes should be of approximately the same discriminability, (4) predicts that *a*, *b*, should be more discriminable than *c*, *d* and (2) predicts the reverse. Thus, the experiment supports theory A *if* (2) is the correct interpretation, which it clearly is not, according to Sutherland's own writings. The other three interpretations yield predictions at variance with the outcome of the experiment; theory C makes a correct prediction.

The writer attempted to devise figures which would yield a more straightforward test between the two theories, but was unsuccessful. The difficulty resides not merely in the fact that several different interpretations of theory A are possible, but also in the nebulous way each one of the interpretations has been put forward. These different interpretations in general do not yield identical predictions, which makes it virtually impossible to test the theory—or theories—with any degree of rigour.

It may be noted that the experiment is biased against theory C in two ways: (i) All group II animals should have had some tendency to treat their positive shape as negative, on either theory, but this is not true of group I; (ii) "surface" shapes were used rather than "contour" shapes, which may introduce a bias in favour of theory A. This assertion cannot be made with confidence, however, since it is not certain whether theory A is supposed to be a theory of surface or contour discrimination. This point has been elaborated in detail elsewhere and the added problems it poses for theory A discussed (Dodwell, 1961 *b*).

CONCLUSION

In confirming a prediction from theory C the presently reported finding tends to support that theory. It is not clear, however, that it really refutes theory A. This is so because it is difficult to determine just what theory A would predict. Also, of course, it is entirely possible that theory A in one version or another applies to the Octopus but not to the rat.

This research was supported by grants from the National Research Council of Canada and the Committee on Scientific Research of Queen's University, whose assistance is gratefully acknowledged.

REFERENCES

DEUTSCH, J. A. (1960). The plexiform zone and shape recognition in the Octopus. *Nature*, **185**, 443–6.

DODWELL, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.*, **48**, 221–9.

DODWELL, P. C. (1960). Discrimination of small shapes by the rat. *Quart. J. exp. Psychol.*, **12**, 237–41.

DODWELL, P. C. (1961 *a*). Visual orientation preferences in the rat. *Quart. J. exp. Psychol.*, **13**, 40–7.

DODWELL, P. C. (1961 *b*). Facts, and theories of shape discrimination. *Nature*, **191**, 578–81.

EDWARDS, A. L. (1960). *Experimental Design in Psychological Research.* New York.

SUTHERLAND, N. S. (1957). Visual discrimination of orientation and shape by the Octopus. *Nature*, **179**, 11–3.

SUTHERLAND, N. S. (1958 *a*). Visual discrimination of shape by Octopus: squares and triangles. *Quart. J. exp. Psychol.*, **10**, 40–7.

SUTHERLAND, N. S. (1958 *b*). Visual discrimination of the orientation of rectangles by *octopus vulgaris* Lamarck. *J. comp. physiol. Psychol.*, **51**, 452–8.

SUTHERLAND, N. S. (1960 *a*). The visual discrimination of shape by Octopus: squares and rectangles. *J. comp. physiol. Psychol.*, **53**, 95–104.

SUTHERLAND, N. S. (1960 *b*). Theories of shape discrimination in Octopus. *Nature*, **186**, 840–4.

PRACTICE, FAMILIARITY AND READING RATE FOR WORDS AND NONSENSE SYLLABLES

BY

R. CONRAD

From the Medical Research Council Applied Psychology Research Unit, Cambridge

Two experiments were carried out to test a number of hypotheses regarding the effects of practice and S–R familiarity on choice reaction times. In Experiment 1 subjects read lists of nonsense syllables of varying levels of association value and differing degrees of choice. It was found that the slope of reading time per syllable was linear with respect to the log of the number of choices, but that the slope was steeper for syllables of lower association value. In Experiment 2, 4- and 32-choice lists of the same average association value nonsense syllables were read on three successive days, as were lists of common three-letter words. The words showed zero slope against number of choices, and the slope for the nonsense syllables decreased with practice. These results are considered to support recent modifications to views on the role of information theory in psychology.

Recent discussions (Welford, 1960; Leonard, 1961) have drawn attention to the effects of practice in multi-choice reaction time experiments. The latter author in particular has cited a number of experiments which show that the classical relationship, whereby RT increases linearly as the number of alternative choices successively doubles, disappears with sufficient practice. Both authors discuss the implications of this effect for human performance models based on information theory. Since at the moment there are no known cases where extended practice has not reduced the difference in RT between two-choice and more than two-choice situations, the evidence that this is a genuinely natural effect is strong.

Another series of experiments cited by Leonard show that as the correspondence between stimulus and response departs from what appears *a priori* to be a natural one, so the slope of RT against number of choices becomes steeper. It is not then, surprising to find that the effects of S–R incompatibility can be reduced by practice (Morin and Grant, 1955). (One hesitates to refer to Fitts and Seeger (1953) in support of this statement, since these authors summarize their experiments by saying that "Differences in reaction time (between different S–R compatibility arrangements) . . . persisted over the 32 days." Yet in fact their data show that the difference between their most and least compatible ensembles was about 0·05 sec. on Day 32 compared with 0·46 sec. on Day 1.)

It therefore appears that effective S–R compatibility can be improved either by changing the relationship between stimulus mode and response mode, or by extended practice of any particular relationship. However the establishment of the relationship between S–R compatibility and reaction time is to some extent circular. Conventionally, two (e.g.) different display-control arrangements are set up, and tests show two different sets of reaction times. From this it is inferred that one arrangement is more compatible than the other. In other words, compatibility is defined in terms of reaction time. One wonders in fact, what conclusion could be drawn in this type of situation if no difference in reaction time was found. Would it be that reaction time is independent of S–R compatibility; or would it be that the S–R compatability of the two arrangements was the same?

One way of avoiding this difficulty would be to use stimuli and responses, the relationship between which could be differentiated in performance terms without recourse to the reaction time. A close approximation to this ideal is possible if the stimuli are nonsense syllables, and the response is a verbal read-out of the syllables presented. For instance, the response (read-out) to the syllable CAT is relatively unequivocal. The response to QEX is less certain. Conventional measures of of syllable difference such as meaningfulness or association value can then be used as measures of doubt about the stimulus-response relationship. In the following experiments the latter measure was accepted as an index of S–R compatibility not based on the time to react to syllables. It is clear that this concept of S–R compatibility is closely related to that of population stereotypes (Fitts and Deininger, 1954). In this sense, ordinary common English words would be regarded as being of higher association value (S–R compatibility) than even 100 per cent. nonsense syllables.

Generalizing then from the light-stimulus/key-response RT experiment to a situation where subjects read out lists of nonsense syllables or similarly constructed ordinary words, a number of predictions can be made.

(*a*) The reading rate for nonsense syllables will depend on the number of different syllables in the test set. The rate will be lower when there are more alternatives to choose from.

(*b*) Given this relationship, the slope will be steeper for syllables of lower association value.

(*c*) For a given level of association value, the slope of the function of reading rate against number of alternatives will decrease with practice.

(*d*) If familiar words are used instead of nonsense syllables, there will be no effect of practice as with (*c*) and as an extension of (*b*) there will be zero slope.

These predictions were tested in the course of two experiments.

EXPERIMENT 1

Method

Twelve sheets of consonant-vowel-consonant nonsense syllables were prepared in lower case type. Each sheet contained 320 syllables, in rows across the page, with wide spaces between syllables and between rows.

The sheets comprised either 4, 8, 16 or 32 different syllables occurring randomly, with the constraint that runs longer than two were broken up, and that every syllable appeared equally often. Each degree of choice was thus represented by three lists of syllables which differed in average association value determined by reference to Glaze (1928). The 12 experimental conditions therefore consisted of four degrees of choice and three levels of association value, the average values of the latter being 15, 46 and 95 per cent. An average was needed because selection of syllables from Glaze's list was restricted in a number of ways. No syllable occurred in more than one list. In any one list initial consonants, and vowels, were varied as much as possible.

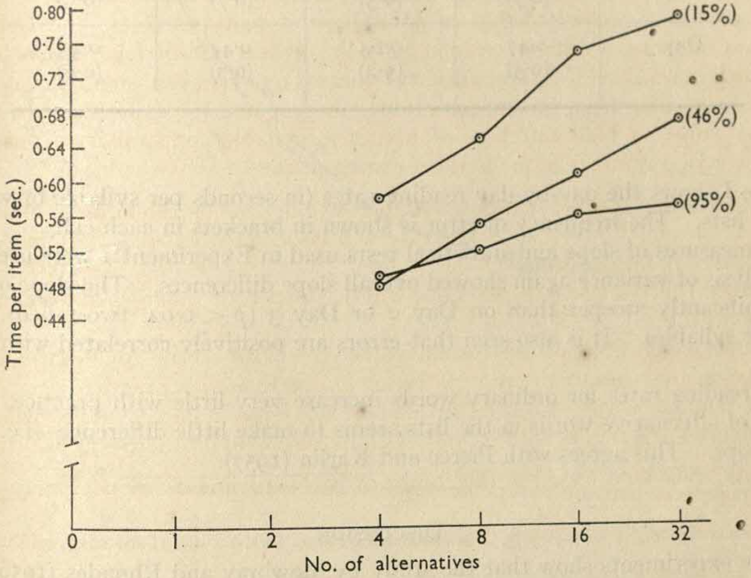
The subjects, who were 14 naval ratings, were given the lists one at a time, and told to read out the syllables as fast and accurately as they could. The lists were given in a different random order to each subject, four per day, on each of three successive days.

Reading time was measured by stop-watch, and the number of errors made were noted.

Results

Figure 1 shows the general trend of the results. It is fairly obvious that with increasing association value of the nonsense syllables, the slope of reading rate against number of alternatives becomes flatter. From the appearance of the slopes, it seems reasonable to take, as a measure of slope, the reading rate difference between four and 32 alternatives. The larger this difference the steeper the slope would be. A Friedman two-way analysis of variance by ranks shows overall differences among the three slopes. The Wilcoxon matched-pairs signed-ranks test, indicates that the 15 and 46 per cent. slopes are both steeper than the 95 per cent. slope at better than the 0·01 significance level (two-tailed). The third comparison yields no difference in slope. Remembering that Glaze's list is undoubtedly subject to a sampling error, and that we used relatively few syllables from it, the failure of the third comparison could be accounted for in this way. This weakness in the design was corrected in the second experiment.

FIGURE 1



Association value of nonsense syllables, choice, and reading time.

There are obvious difficulties about scoring errors of reading nonsense syllables, since pronunciation is inevitably arbitrary. What is important if speed is the main criterion considered, is that error rate should not be high, and that speed and accuracy should not be negatively correlated. Overall, we found error rates of about 3 per cent. with a clear indication that the few errors which did occur were positively correlated with speed.

EXPERIMENT 2

Method

Eight different naval ratings read through four lists prepared as before. They read them again the next day, and once more on the third day. Every subject had the same four lists each day, but in a different order, and this order was also varied amongst subjects.

The lists again contained 320 syllables each composed as follows:—

List A.—Four different 20 per cent. association value nonsense syllables.
B.—Thirty-two ditto.
C.—Four consonant-vowel-consonant common English words taken from the most frequent words in Thorndike and Lorge (1944).
D.—Thirty-two ditto.

To minimize the possibility that the four-item vocabularies were by chance easier than those of 32 items, the four different nonsense syllables and four different words in the A and C lists respectively, were varied for each subject.

TABLE I

READING RATE (SEC. PER ITEM) AND ERRORS FOR WORDS AND NONSENSE SYLLABLES

	List A	*List B*	*List C*	*List D*
Day 1	0·59 (1·8)	0·86 (7·7)	0·45 (0·9)	0·46 (0·7)
Day 2	0·51 (0·7)	0·69 (6·5)	0·43 (0·7)	0·43 (0·9)
Day 3	0·47 (0·5)	0·64 (5·2)	0·42 (0·5)	0·43 (0·6)

Results

Table I shows the day-by-day reading rates (in seconds per syllable or word) for the four lists. The frequency of error is shown in brackets in each cell.

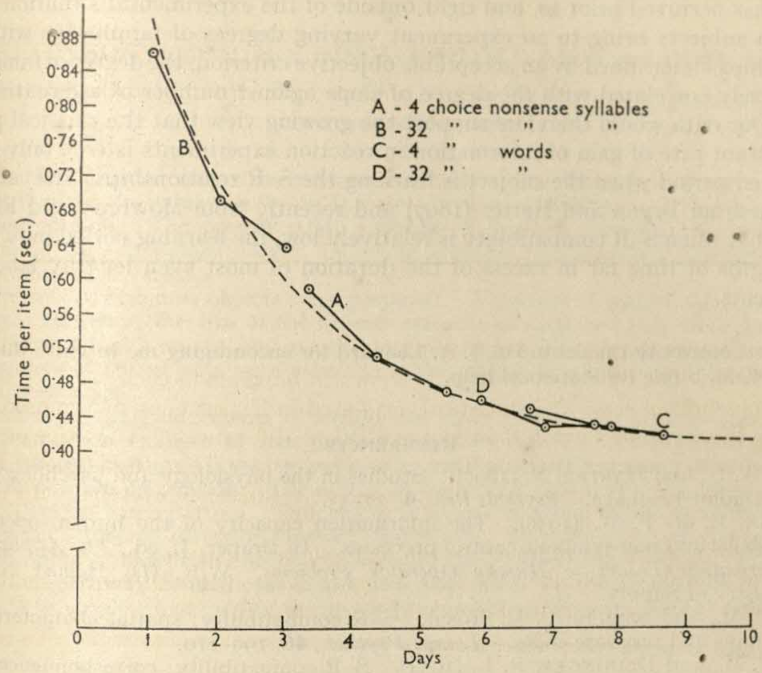
The measures of slope and statistical tests used in Experiment 1 were used again. The analysis of variance again showed overall slope differences. The slope on Day 1 was significantly steeper than on Day 2 or Day 3 ($p < 0{\cdot}02$, two-tailed), for the nonsense syllables. It is also seen that errors are positively correlated with reading rate.

The reading rates for ordinary words increase very little with practice, and the number of alternative words in the lists seems to make little difference—i.e. there is a zero slope. This agrees with Pierce and Karlin (1957).

DISCUSSION

These experiments show that the study by Mowbray and Rhoades (1959) on the effect of extended practice on choice reaction time, is in fact closely related to that of Crossman (1956); Crossman showing that a symbolic S–R relationship gave a steeper slope of RT against choice than a non-symbolic relationship. By implication this point had already been made in Leonard's experiment (1959); his response to a tactual stimulus might reasonably be taken as more highly practiced than to a light stimulus. Equally, directly stimulating the finger from which response is required, can be taken as a more compatible arrangement than requiring subjects to interpret even an apparently simple light display. Our two experiments together make the argument more explicit. Experiment 1 is analogous to that of Crossman, in that the S–R situation is made increasingly familiar, leading to a flatter slope. But whereas in conventionally varying S–R compatibility, Crossman assumed that his two modes were of different difficulty, an assumption confirmed by his results, in our case, the difficulty difference could be directly written into the experimental conditions. But in examples where S–R compatibility is varied, the differing degrees of familiarity with the S–R relationships do not develop during test sessions. The fact is of course that the experimental conditions use the implied practice that has occurred differentially outside of the experimental situation. In this sense, Leonard's tactual response represents yet a further stage in the Crossman situation. Mowbray and Rhoades reach this stage by making subjects explictly, inside the experimental situation, go through the practice drill. There can be little doubt that had Experiment 1 carried another four conditions using ordinary words, the result would have been similar to that found in Experiment 2—namely zero slope. This procedure was discarded merely for the sake of experimental economy.

FIGURE 2



The data of Table I fitted to an exponential hypothetical practice curve.

Experiment 2 reaches the same conclusion by the other route. Here we take S–R combinations which we know through external criteria to have been relatively infrequently practiced up to the time when the subject arrived at the experiment. We show that practice alone during the experiment reduces the slope of reading time against number of alternatives. The conditions (C and D) using words, represent a short cut to the end of the practice/familiarity axis. This contention is supported by the re-arrangement of the data in Figure 2. This shows the effect of fitting an exponential imaginary practice curve to the results shown in Table I; the curve is calculated to reach zero RT at infinity. The method of arriving at Figure 2 requires explanation. A suitable curve was calculated to represent a hypothetical course of practice. The ordinate scale is determined by the experimental data (Table I), the scale of the abscissa is arbitrary equal time intervals. The 3-day plot for each experimental condition, correctly sited on the vertical axis, was adjusted horizontally to give the best fit by eye to the hypothetical practice curve. By trial and error a practice curve was found to which all four sets of experimental data could satisfactorily be fitted. The result shown suggests that all the points could lie on a single time continuum which would be represented by successive days of practice. Thus one feels more justified in extrapolating the changes of slope indicated by Table I from the nonsense syllable material to the words.

Summarizing, we have argued that the time to respond to a series of stimuli depends on the familiarity of the S–R relationship. When this is new to subjects, the mean response time is a function, with a positive slope, of the number of alternatives. The relationship becomes more familiar through practice; in which case the slope trends to zero. Practice can be shown to be effective in this way when it is carried out as part of the experimental situation. Equally it can be effective even when it has occurred prior to, and right outside of the experimental situation.

When subjects bring to an experiment varying degrees of familiarity with S–R relationships, determined by an acceptable objective criterion, the degree of familiarity is negatively correlated with the degree of slope against number of alternative S–R units. Our data would therefore support the growing view that the classical picture of a constant rate of gain of information in reaction experiments is true only during the limited period when the subject is learning the S–R relationships. Nevertheless it is clear from Bryan and Harter (1897) and recently from Mowbray and Rhoades (1959), that when S–R compatibility is relatively low, the learning period may extend over lengths of time far in excess of the duration of most even lengthy laboratory experiments.

I must record my thanks to Dr. J. A. Leonard for encouraging me to write this paper and to Dr. M. Stone for statistical help.

References

Bryan, W. L., and Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychol. Rev.*, **4**, 27–53.

Crossman, E. R. F. W. (1956). The information capacity of the human operator in symbolic and non-symbolic control processes. In Draper, J., ed., *The Application of Information Theory to Human Operator Problems*. W.R. (D) Report No. 2/56 Ministry of Supply.

Fitts, P. M., and Seeger, C. M. (1953). S–R compatibility: spatial characteristics of stimulus and response codes. *J. exp. Psychol.*, **46**, 199–210.

Fitts, P. M., and Deininger, R. L. (1954). S–R compatibility: correspondence among paired elements within stimulus and response codes. *J. exp. Psychol.*, **48**, 483–92.

Glaze, J. A. (1928). The association value of nonsense syllables. *J. genet. Psychol.*, **35**, 255–67.

Leonard, J. A. (1959). Tactual choice reactions: I. *Quart. J. exp. Psychol.*, **11**, 76–83.

Leonard, J. A. (1961). Choice reaction time experiments and information theory. In Cherry, C., ed., *Proceedings of the 4th London Symposium on Information Theory*. London.

Morin, R. E., and Grant, D. A. (1955). Learning and performance on a key-pressing task as a function of the degree of spatial stimulus-response correspondence. *J. exp. Psychol.*, **49**, 39–47.

Mowbray, G. H., and Rhoades, M. V. (1959). On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

Pierce, J. R., and Karlin, J. E. (1957). Reading rates and the information rate of a human channel. *Bell System Tech. J.*, **36**, 497–516.

Thorndike, E. L., and Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York.

Welford, A. T. (1960). The measurement of sensory-motor performance: survey and reappraisal of 12 years' progress. *Ergonomics*, **3**, 189–230.

SIZE-WEIGHT INTERACTION IN JUDGEMENTS OF COMPOUND STIMULI

BY

W. S. ANTHONY[1], D. H. HOLDING[2], JUDITH S. LION and W. SLUCKIN[3]

From the Department of Psychology and the Nuffield Department of Industrial Health
University of Durham

Recent work in the fields of multidimensional information transmission and perceptual accentuation due to value has led to theories concerning unidimensional judgements of compound stimuli. The present study attempts to explore these theories, adding the dimension of size to judged weight, following negative results with judgements of size when brightness is varied.

Three sets of stimulus objects were prepared. Variation of weight was the same in each set. However, the size of the objects was the same in one set, while in another set size increased with weight, and in a third set size and weight were negatively correlated. Three groups of 16 subjects were asked to make replicated absolute judgements of the weights in each set.

It appeared that judgements of weight are more affected by density when density varies than when changes of size produce constant density. As predicted, contrast between successive stimuli is greatest when size and weight are negatively correlated, and least with positive correlation. The extent of the judged scales, and the means of these scales, follow the same pattern. Differences in discriminability between the three series, whilst in the expected direction, are not significant. The hypothesis that these effects would be more apparent in the psychologically linked dimensions of size and weight, than in the relatively neutral case of size and brightness, was largely confirmed. These findings both amplify and modify the previous suggestions concerning compound stimuli.

INTRODUCTION

The present experiment follows on work by Sluckin and Holding (1959) and relates particularly to the analysis of perceptual accentuation by Tajfel (1957, 1959). In examining one-dimensional judgements of stimuli varying in two dimensions, the experiment throws light on the size-weight illusion (*a*) as such, and (*b*) as relating to other multidimensional effects, and to the value-size and value-weight effects shown by Messenger (1902), Bruner and Goodman (1947) and others. It will be convenient to introduce the two aspects of the work under separate headings.

The size-weight illusion as such

The size-weight illusion is commonly demonstrated and quantified by obtaining a point of subjective equality of weights between objects of different size; of two objects, the smaller has to be physically less heavy than the larger, if they are to seem equally heavy to the subject; in other words, the weight of the smaller object is overestimated relative to the weight of the larger object. The illusion may be reduced in some conditions (e.g. Huang, 1945*a*), and reversed at one extreme (Howard, 1954); but in this paper we shall not be concerned with these special conditions.

The illusion, demonstrated in the normal way, may be regarded as only one of several possible effects resulting from a more general interaction of the dimensions of size and weight in subjective judgements. Other effects resulting from the general

[1] Now at Department of Psychology, University of Auckland, New Zealand.
[2] Now at Department of Psychology, University of Leeds.
[3] Now at Department of Psychology, University of Leicester.

mechanism of perception should be demonstrable in cases where the illusion, as normally defined (or demonstrated), does not appear. One such case is where absolute judgements are made of the weights of a series of stimuli which are compound, in the sense that they also vary in the size dimension.

In the present experiment we present the subject with a set of objects, one by one, and we ask him to estimate the weight of each object. From our knowledge of the size-weight illusion we expect that the smaller an object is the heavier it will be judged. Although the subject is making an absolute judgement we assume that as each object is presented to him he is comparing the object with the previous one. This assumption formed the basis of a measurement which we call "subjective contrast." This is based on the differences between judgements to successive stimulus presentations and is described more fully in the Results section.

In one set of objects, the Neutral set, size remains constant. In another set, the Plus set, size and weight co-vary positively, and density remains constant. For this set, as the subject weighs one object after another, an object which is actually lighter than the preceding one is also smaller, so we expect that the difference of size will produce relatively small differences between the judgements, and that the subjective contrast will be small. Similarly an object which is heavier than the preceding one should be judged heavier to a lesser degree than if size were constant; thus throughout the sequence of presentations the differences between successive judgements should be small relative to the Neutral condition where size is constant. In the Minus set of objects size and weight co-vary negatively (the heavier objects are smaller than the lighter ones). Here following a similar line of reasoning we expect that the differences between successive judgements will be large relative to the Neutral condition. For the subject, contrast between the stimuli composing the series will be high. In this experiment therefore we expect to find that subjective contrast will be least, medium and most for the Plus, Neutral and Minus sets respectively.

If this prediction is confirmed, then we have demonstrated in this situation the analogue of the size-weight illusion, and we may proceed to test hypotheses concerning it.

The size-weight illusion as relating to other multidimensional effects

Many experimenters have demonstrated the influence of value on judgements of physical magnitude, especially size and weight. Typically, when value and size (or weight) are co-varied in a series of stimulus-objects, and when the subject's task is to estimate size (or weight), the objects are judged more different in size (or weight) than when value is not present in the series of objects. This generalization is proposed by Tajfel (1957, 1959) (also see McCurdy, 1956) and it is supported by results on the apparent sizes of coins (Bruner and Goodman, 1947; Bruner and Rodrigues, 1953; Carter and Schooler, 1949), and by results on apparent lengths and weights (Dukes and Bevan, 1952; Tajfel, 1957) where correlative value was experimentally induced. Here we are not concerned with subjective contrast but with what we shall call scale extent (difference between the mean judgements of the lightest and heaviest weights), a measure of range.

In the work cited, the presence of value increased the scale extent. Tajfel (1957) suggests that this effect may be related to the increased discrimination obtained with purely physical dimensions by Eriksen and Hake (1955 *b*) and others. Eriksen and Hake found that stimuli in which size, hue, and brightness co-varied were better discriminated than stimuli in which only one dimension varied, or only two co-varied. Their subjects were set to identify stimulus objects, and presumably felt free to use all available cues.

In Sluckin and Holding's (1959) experiments, brightness and size were both varied but, as in the value experiments, the subjects were set to judge only one dimension (size). The effect of the co-varying brightness was on the whole insignificant. The experimenters concluded that the generalization (suggested by Miller, 1956) that multidimensional stimuli yield increased transmission of information, may only hold when the subject is set to respond multidimensionally, or alternatively when combinations of dimensions are used which are psychologically linked. Size and weight are one such combination, since all the work on the size-weight illusion shows that even when the subject is restricting his attention to weight, his judgements are affected by difference in size. Thus size and weight, unlike size and brightness, may provide results analogous to the value effects, and in conditions more analogous than the Eriksen and Hake situation. An analogy between the value effects and the size-weight illusion has previously been vaguely implied by Ansbacher (1937), and has been developed by Tajfel (1957).

In our present experiment, we pursue this analogy in more detail. Although subjective contrast and scale extent are different measures, a homogeneous increase in subjective contrast must be accompanied by an increase in scale extent, so we expect to find that scale extent will be least, medium and greatest for the Plus, Neutral and Minus sets respectively.

The influence of the size-weight illusion on discrimination and mean response will be dealt with in the Results and Discussion.

METHOD

Subjects

The subjects were 48 members of, or workers for, Durham University or their friends. None had taken a course in general or experimental psychology. They were between 16 and 60 years in age.

Apparatus

Three sets of objects were prepared, 10 objects in each set, making 30 objects in all. The objects were cylindrical, the height of each being equal to the diameter; they were painted a uniform red.

The sets did not differ in terms of weight; in each set, the lightest object, object 1, was 100 gm., the next lightest was 110 gm., the next was 121 gm., and so on by increases of 10 per cent, the heaviest object, object 10, weighing about 236 gm. The objects were accurate to within 0·3 per cent. of the ideal weights. It is to be noted that object 6 in each set weighed about 161 gm. (this value was given to the subject at the beginning of the experiment, as explained below).

The sets differed in terms of size. In the Plus set, the objects were as big in cubic centimetres as they were heavy in grammes; that is, object 1 was 100 c.c. in volume, object 2 was 110 c.c., and so on. In the Minus set, the sizes were reversed; that is, object 10 was 100 c.c. in volume, object 9 was 110 c.c., and so on. In the Neutral set, the size of all the objects was 150 c.c.

All the objects were marked on one end of the cylinder in white paint, "+1," "−7," "N10," etc., for identification by the experimenter. Subjects, by instruction, held the objects "bottom" downwards (see below).

Procedure

Only one set of objects was used in connection with any given subject; the subjects were divided into three equal groups of 16, the Plus, Neutral, and Minus groups, according to the set of objects used.

Subjects were taken individually. In the experimental situation, the subject and experimenter sat opposite each other, across a small table. To the right hand of the experimenter were placed the set of objects to be used and a sheet of paper which indicated the order of presentation of objects, and on which the subject's responses were recorded by the experimenter. The set of objects and the sheet of paper were excluded from the subject's view by a screen.

When the subject and experimenter had sat down, the experimenter took object 6 from behind the screen and placed it in view of the subject, and spoke the following instructions:

"When I give you an object, I want you to weight it in your hand; tell me how heavy you think it is, and give it back to me. The first object weighs one hundred and sixty-one grammes, one sixty one, and I want you to give all your answers like that, in grammes. Now, all the objects are marked on the bottom, so I want you to be careful to handle the objects without turning them over, otherwise it will spoil the experiment. Here is the first object: it weighs one hundred and sixty-one grammes. Lift it up and down a little,

TABLE I

Random Orders for the First Five Subjects in Each Group

(Orders for subsequent subjects follow those given in the table, viz. sixth subject as first, seventh as second, etc.)

Subject	1	2	3	4	5
Weights	121	177	100	236	177
	146	214	214	177	110
	236	100	195	110	195
	195	161	110	161	146
	110	121	133	100	133
	100	236	161	121	161
	133	110	177	195	100
	161	195	236	214	236
	177	146	121	133	214
	214	133	146	146	121
	133	161	236	214	133
	110	133	110	161	110
	146	121	161	121	177
	100	195	121	236	236
	121	146	100	100	195
	177	214	177	146	146
	214	236	214	195	100
	161	100	146	110	121
	195	110	195	133	214
	236	177	133	177	161
	236	236	195	195	161
	133	214	133	146	100
	214	195	121	177	236
	146	161	236	110	110
	161	133	146	100	146
	177	110	161	121	177
	195	146	214	133	195
	121	121	110	214	121
	110	177	100	161	133
	100	100	177	236	214
	100	161	195	100	214
	195	133	100	195	146
	121	195	110	133	161
	177	214	121	121	100
	161	177	214	110	236
	133	236	236	161	121
	110	100	161	177	133
	146	146	133	146	195
	214	110	177	236	177
	236	121	146	214	110
Objective contrast	37·2	48·1	54·5	51·8	49·1

to try to get an idea of its weight, and then hand it back to me." While the third sentence was spoken, the experimenter lifted object 6 up and down in his/her hand, the bottom (the marked end) of the object resting on the palm, to suggest a manner of weighing the objects.

When the subject had handed back object 6, the experimenter proceeded with the main part of the experiment, which consisted of forty presentations according to a predetermined order (see Table I). This order was composed of four sub-sequences: the ten objects were presented, one by one, in a random order (the orders had been obtained from tables of random numbers); then they were presented again, in another random order; and then for a third, and then for a fourth time, each time in a different random order. There was no special break between the sub-sequences, or indication of them to the subject. Object 6, like the other objects, was presented four times, with no indication that it had also been the object given initially as a standard.

To each member of the Plus group, there corresponded a member of the Neutral group and a member of the Minus group, each of these three subjects being given the same order of presentation, by weight.

The procedure usually took about 15 min. for each subject. Subjects were not restricted to the use of one hand, or to any specific manner of weighing the objects, except that the bottom was kept downwards.

RESULTS

The results are presented in Tables II, III, IV and in Figure 1.

Subjective contrast

First, the objective contrast was calculated for each "random order" (see Table I). This was done by summing the objective differences between successively presented weights. The first weight in the series was deducted from the initial standard of 161 gm., making 40 differences in all.

A measure of subjective contrast was then obtained for each subject by summing the differences between the subject's judgements counting differences in the wrong direction as negative. (For example, if a subject said "150" in response to object 1 and then was given object 3 and said "100," this was counted as a difference of minus 50.) The sum of these differences was then divided by the sum of the objective differences. This quotient was then multiplied by 100 to give a percentage score.

Examining Table II we see that our prediction concerning subjective contrast in the three groups is confirmed. The null hypothesis that the groups do not differ in subjective contrast may be rejected at the 0·001 level, by the Kruskal-Wallis test. Applying the one-tailed Mann-Whitney U-test, we find that each group is significantly different from each other, at the 0·05 level or better.

TABLE II

SUBJECTIVE SCORES, IN FOUR VARIABLES, FOR THE THREE GROUPS EXPRESSED AS PERCENTAGES OF CORRESPONDING OBJECTIVE SCORES. (FULL EXPLANATION SEE TEXT.)

	Means for each group			
Variable	*Plus*	*Neutral*	*Minus*	*Limits of p by Kruskal-Wallis test*
Subjective contrast	63·4	109·5	174·2	$0·001 > p$
Scale extent ..	64·8	104·5	162·3	$0·001 > p$
Mean response ..	99·4	117·4	122·7	$0·05 > p > 0·02$
Discrimination ..	63·7	70·9	72·3	$0·10 > p > 0·05$

Scale extent

For each subject, we subtract the mean response to object 1 from the mean response to object 10. The objective difference is 136 (135·8) gm. Table II gives the mean extent of scale, for the three groups, as a percentage of this objective extent. The Kruskal-Wallis test again rejects the null hypothesis at the 0·001 level. The differences between the groups are as expected, although the difference between the Plus and Neutral groups is not significant by one-tailed U-test.

Table III shows the means and standard deviations of the judgements from each group to the ten weights, together with the physical values of the weights.

In Figure 1 mean response magnitudes are plotted against the ten stimulus magnitudes. The scales of judgement are shown as the projections of the empirical points on to lines parallel with the response axis. The physical scale is shown alongside for comparison.

TABLE III

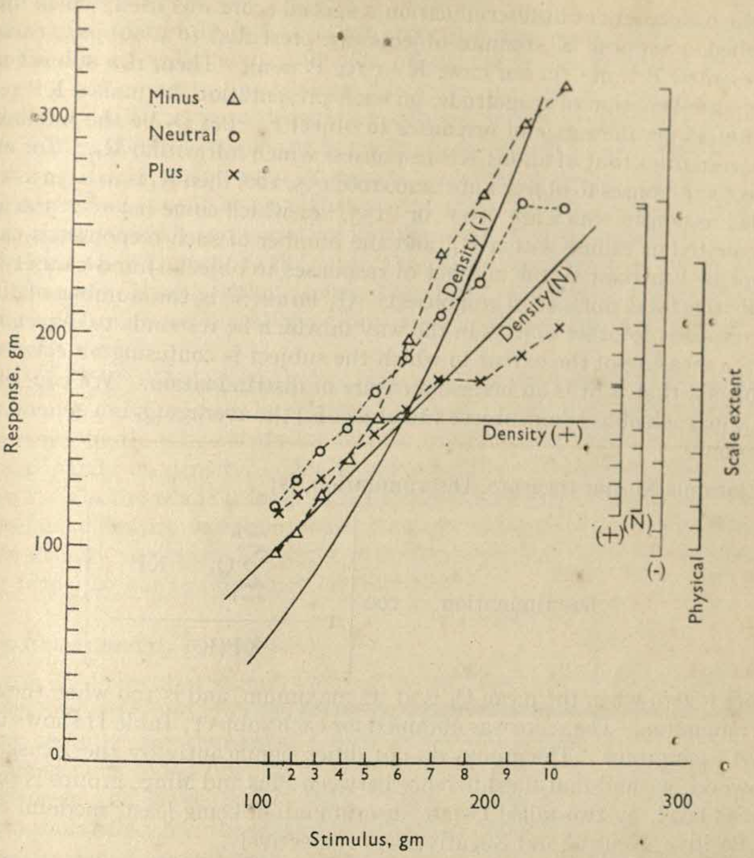
MEANS AND STANDARD DEVIATIONS OF THE JUDGEMENTS FROM EACH GROUP TO THE TEN WEIGHTS

Physical weight in gm.	*Average judgement (and standard deviation) of weight in gm.*					
	Plus		*Neutral*		*Minus*	
	Mean	σ	*Mean*	σ	*Mean*	σ
100·0	116·7	31·1	118·6	32·0	97·6	37·7
110·0	123·4	30·7	131·9	31·6	107·6	40·6
121·0	131·9	29·6	145·8	24·0	124·9	38·1
133·1	139·6	24·3	156·2	26·5	142·1	36·1
146·4	152·2	26·0	173·2	28·9	161·9	43·7
161·0	166·3	27·7	191·8	37·7	196·9	58·0
177·1	179·6	40·0	209·0	60·2	238·1	63·5
194·8	177·8	43·4	224·9	57·9	265·6	87·5
214·3	191·6	38·0	262·1	81·0	303·1	102·6
235·7	204·7	38·3	260·4	72·9	317·9	91·6

Mean response

It will be noticed that all judgemental scales, including the Neutral series, are displaced relative to the corresponding physical scales. The extent of this shift, which represents an effective change in adaptation level, has been measured by averaging the mean response levels of the subjects in each group. These group means, expressed as percentages of the mean stimulus magnitude (159·4 gm.) are shown in Table II. The Kruskal-Wallis test shows the differences between all groups to be significant at the 0·05 level. Applying the two-tailed U-test to the individual comparisons, the difference between the Neutral and Minus groups is not significant, and the Plus-Neutral and Plus-Minus differences are both significant at the 0·05 level. Owing to the fact that a subjective scale can extend more away

FIGURE I



Mean response magnitudes plotted against the ten stimulus magnitudes together with density variations for the three groups.

from zero than towards it, it was expected that scale extent and mean response would be positively correlated. In order to confirm that these relationships hold good *within* the groups, Spearman's rho correlations between scale extension and mean response level were calculated from the data obtained under each condition. These are shown in Table IV. As expected all the correlations are positive; although on a two-tailed test only the value for the Neutral condition is significant ($p < 0{\cdot}02$).

TABLE IV

Correlated variables	*Spearman's rho, for each group*		
	Plus	*Neutral*	*Minus*
Scale extent and mean response . .	0·341	0·624 ($0{\cdot}02 > p$)	0·514 ($0{\cdot}10 > p > 0{\cdot}05$)
Scale extent and discrimination . .	0·076	0·402	0·271

Correlations: all are positive; by two-tailed test, one is significant, one approaches significance

Discrimination

For the measurement of discrimination a special score was used, which must now be described. Suppose K stimulus-objects are presented to a subject, each object being presented P times (in our case, $K = 10$, $P = 4$). Then, if a subject makes a response, an estimation of magnitude, on each presentation, he makes KP responses in all. Let R_r be the range of responses to object r. Let Q_r be the number of the subject's responses (out of all his KP responses) which fall within R_r. For example, if a subject's responses to object 1 are: 120, 100, 155, 100, then R_1 is written as 100–155. Any of his responses which are "100" or "155," or which come between 100 and 155, will be counted as falling within R_1, and the number of such responses is called Q_1. Q_1 cannot be less than P (the number of responses to object 1) and cannot be more than KP (the total number of responses). Q_1 *minus* P is the number of times the subject responds to other objects in the way in which he responds to object 1, and is therefore a measure of the extent to which the subject is confusing object 1 with the other objects; that is, it is an inverse measure of discrimination. We can, of course, obtain Q_r for each object from object 1 to object K; the average Q_r is a general measure of confusion.

The formula for our measure, Discrimination, is:

$$\text{Discrimination} = 100\left[1 - \frac{\sum_{r=1}^{r=K} Q_r - KP}{KP(K-1)}\right]$$

This score is zero when the mean Q_r is at its maximum, and is 100 when the mean Q_r is at its minimum. The score was obtained for each subject; Table II shows the mean result for each group. The groups do not differ significantly by the Kruskal-Wallis test; however, we find that the difference between Plus and Minus groups is significant at the 0·05 level, by two-tailed U-test, discrimination being least, medium and most for the Positive, Neutral and Negative sets respectively.

Variations in judgement compared with variations of weight and density

In order to see how far the judgements are influenced by density it is necessary to compare the obtained curves with the theoretical curves for density. In the case of the Neutral group, density and weight changes coincide. For the other two groups, density curves have been calculated using the density of the standard object (No. 6) as a reference point (see Fig. 1).

The judgements for the Neutral group appear to be parallel to the appropriate weight and density curve. In the case of the Minus group the lighter judgements follow the density curve rather closely with some divergence towards weight in objects 8, 9 and 10. The same objects in the Plus curve, which otherwise follow weight fairly accurately, are inclined towards the appropriate density curve. Since, however, this inclination to the right occurs to some extent in all groups, it may be better interpreted as implying that the underlying scale of increment is not logarithmic rather than as providing weight density information.

DISCUSSION

(1) *Size-weight illusion*

As pointed out in the introduction the prediction that subjective contrast will be least, medium and most for the Plus, Neutral and Minus sets respectively, derives logically from the usual demonstration of size-weight illusion. This leads us to examine the role of density in size-weight judgements.

Ryan (1940) writes "in fact *O* is not reporting weight at all but *specific gravity* and the illusion is really not an illusion at all. *O* is simply unable to report on weight being influenced by the striking differences in density involved." This view which we may call the pure density hypothesis, seems to overstate the case, for apparently the subject's judgement is generally *in between* what would be correct in terms of weight and what would be correct in terms of *specific gravity*. Following Ryan's theory one would expect that in the Plus set (in which density remains constant) there would be no variation in judgement of different weights, i.e. each of the 10 weights would appear to the subject to be the same weight. As we know, this did not happen. Ryan's theory must, therefore, be rejected although in certain circumstances (such as in our Minus group in the middle range of weights) it may be substantially correct.

Thouless (1931), Huang (1945*a*, 1945*b*) and Nyssen and Bourdon (1956) have put forward a "compromise hypothesis" namely, that in making judgements of weights, compromise judgements are in fact being made which are partly based on weight and partly on density. This is supported by our findings and is illustrated in Figure 1. On the whole it looks as if judgements of weight tend to a considerable extent to follow density variations when these are substantial as in the Minus group; but when density variations are slight or non-existent, as in the plus group, there is a strong tendency not to rely much on density cues.

(2) *Multidimensionality of stimuli*

There is no doubt that stimuli varying in more than one dimension are identified more often correctly, and are thus better discriminated from one another, than those that vary in only one dimension. This applies to stimuli which either vary in several dimensions independently or vary along the different dimensions concomitantly (Eriksen and Hake, 1955*b*; Miller, 1956). While two or more dimensions of the stimulus objects vary, the magnitudes may be judged of only a single dimension.

It can be asked what would be the effect of additional dimensions, perfectly correlated with the first, when judgements are made of the first dimension only. Tajfel (1957), quoting Eriksen and Hake (1955*b*), implies that the addition of the second dimension, when judgements of only the first dimension are made, should also result in improved discrimination, provided the second dimension is correlated with the first. Sluckin and Holding (1959) attempted to verify this, using stimuli that varied in size and brightness, size having been the dimension judged. They found, among other things, that discrimination was no better when variations in size were accompanied by concomitant variations in brightness than when brightness remained constant throughout, or varied in some way unrelated to variations in size. They concluded that their experiments had "failed to provide any support for the theory that (*a*) extension of subjective scale, or (*b*) increase in discriminability, will result from the addition of a second physical dimension to a first judged dimension, and that it cannot be asserted that two-dimensional effects do not occur; it can no longer be asserted, however, that they will always occur."

In the present experiment it was possible to test in a new setting the general expectation based on Miller (1956) that the addition of a second correlated dimension, when judgements are made of the first dimension only, should result in improved discrimination. In the present experiment we would expect that in both Plus and Minus sets, where there is variation in size as well as in weight, discrimination would be better than in the Neutral set. Once again the expectation of improvement in discrimination was falsified. Sluckin and Holding (1959) suggested that such two-dimensional effects as occur when one dimension only is judged may, in fact, be illusion effects; that is, the effects occur when dimensions are psychologically linked. The present experiment has shown that "pure" two-dimensional effects, if there are such, can no more be shown to exist in the case of linked dimensions (e.g. size-weight) than in the case of non-linked cues (e.g. size-brightness).

Multidimensional effects (such as enhanced discrimination and extension of scale) seem to occur in certain circumstances when dimensions are psychologically linked, possibly as a result of previous experience of a concomitant variation in the two or more dimensions. This condition is met in the present experiment, so that the unidimensional responses that are given to multidimensional stimuli may be no longer truly unidimensional. Improved discrimination and other effects resulting from the addition of a "dimension" of value, instead of a physical dimension, appear to be similar to the effects arising in situations where subjects have had previous experience of co-variation of dimensions (Lambert, Solomon and Watson, 1949). It is clear from the foregoing that the direction of correlation between the stimulus dimensions is very important when subjects judge only one dimension, although when subjects identify objects, the direction of correlation between stimulus dimensions appears not to matter. Here we are not concerned with direction of correlation in the mathematical sense but wish to indicate whether the second dimension is reinforcing the first one. For example we would say in a series of discs in which the largest ones were the brightest that size and brightness are in the same direction, whereas if the largest discs were the dimmest the correlation would be in the opposite direction. The concept of psychological linkage is necessarily directional; therefore, where such linkage occurs, multidimensional effects are directional, as in our experiment.

(3) *Scale extension and improved discrimination*

We have seen that our prediction that scale extent would be least, medium and greatest for the Plus, Neutral and Minus sets respectively, was confirmed. Tajfel associates increase in scale extent with increase in discrimination and we were interested to examine this relationship in the present experiment. Sluckin and Holding (1959) in the size-brightness experiment found no indication that increase in extent of scale was accompanied by an increase in discrimination. Among non-value experiments (reviewed by Alluisi, 1957), Eriksen and Hake (1955*a*) and again Pollack (1952) showed that an objective increase in stimulus range could produce a slight increase in discrimination. Our problem was whether a subjective increase in stimulus range would give the same result.

Table II shows that discrimination was least, medium and most for the Plus, Neutral and Minus sets respectively. Correlations between discrimination and scale extent whilst being positive were small and lend weak support to Tajfel's theory. This may have been due to the relatively small increase in discrimination as between the Plus, Neutral and Minus sets. There is no ground for doubting that discrimination and scale extent are associated. However, in view of the fundamental importance of this association for qualitative judgements in social perception (Tajfel, 1957, 1959) it would be useful if further psychophysical studies were to be carried out. Such studies should show more precisely the character of the association between scale extent and discrimination in the case of judgements of various purely physical dimensions.

(4) *Overestimation and asymmetrical scale extension*

In the value experiments, a general overestimation (relative to judgements in a control condition) is often reported, rather than greater scale extension (Lambert, Solomon and Watson, 1949; Dukes and Bevan, 1952). In the present experiment this would appear as a shift in mean response and was present although not significant, in Sluckin and Holding's data.

As mentioned earlier it was expected that scale extent and mean response would be positively correlated. This expectation was clearly confirmed even if only one of the three correlations proved significant. It was also found that the mean response was lowest, medium and highest for the Plus, Neutral and Minus sets respectively. In other words, weights in the Minus set were on the average judged heavier than those in the Neutral set, and even more so than those in the Plus set. They were also on the average judged heavier than they were in fact; that is, these weights were on the whole overestimated (and weights in the Neutral set were also overestimated, although not so much as those in the Minus set).

The fact that the average judgements of weight in the three experimental conditions were significantly apart lends strong support to Tajfel's view (1959) that in order to explain overestimation phenomena, one has to consider what happens *within* each series. The present experiments may help to elucidate the ways in which interserial and intraserial effects are related.

REFERENCES

ALLUISI, E. A. (1957). Conditions affecting the amount of information in absolute judgements. *Psychol. Rev.*, **64**, 97–103.

ANSBACHER, H. (1937). Perception of number as affected by the monetary value of the objects. *Arch. Psychol. N.Y.*, No. 215.

BRUNER, J. S., and GOODMAN, C. C. (1947). Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, **42**, 33–44.

BRUNER, J. S., and RODRIGUES, J. S. (1953). Some determinants of apparent size. *J. abnorm. soc. Psychol.*, **48**, 17–24.

CARTER, L. F., and SCHOOLER, K. (1949). Value, need and other factors in perception. *Psychol. Rev.*, **56**, 200–7.

DUKES, W. F., and BEVAN, W. (1952). Accentuation and response variability in the perception of personally relevant objects. *J. Personality*, **20**, 457–65.

ERIKSEN, C. W., and HAKE, H. W. (1955*a*). Absolute judgements as a function of stimulus range and number of stimulus and response categories. *J. exp. Psychol.*, **49**, 323–32.

ERIKSEN, C. W., and HAKE, H. W. (1955*b*). Multidimensional differences and accuracy of discrimination. *J. exp. Psychol.*, **50**, 153–60.

HOWARD, I. P. (1954). An experimental and theoretical analysis of the size-weight illusion. *Durham Res. Rev.*, **1** (No. 5), 47–53.

HUANG, I. (1945*a*). The size-weight illusion in relation to the perceptual constancies. *J. gen. Psychol.*, **33**, 43–63.

HUANG, I. (1945*b*). The size-weight illusion and the "weight-density illusion." *J. gen. Psychol.*, **33**, 65–84.

LAMBERT, W. W., SOLOMON, R. L., and WATSON, P. D. (1949). Reinforcement and extinction as factors in size estimation. *J. exp. Psychol.*, **39**, 637–41.

MCCURDY, H. G. (1956). Coin perception studies and the concept of schemata. *Psychol. Rev.*, **63**, 160–8.

MESSENGER, J. F. (1902). How many dollar bills will equal in weight a five-dollar gold piece? *Science*, **15**, 672–3.

MILLER, G. A. (1956). The magical number seven, plus or minus two. *Psychol. Rev.*, **63**, 81–97.

NYSSEN, R., and BOURDON, J. (1956). A new contribution to the experimental study of the size-weight illusion. *Acta Psychol.*, **12**, 157–73.

POLLACK, I. (1952). The information of elementary auditory displays. *J. acoust. Soc. Amer.*, **24**, 745–9.

RYAN, T. A. (1940). Interrelations of the sensory systems in perception. *Psychol. Bull.*, **37**, 659–98.

SLUCKIN, W., and HOLDING, D. H. (1959). Judgements of compound stimuli. *Durham Res. Rev.*, **2** (No. 10), 303–9.

TAJFEL, H. (1957). Value and the perceptual judgement of magnitude. *Psychol. Rev.*, **64**, 192–204.

TAJFEL, H. (1959). Quantitative judgement in social perception. *Brit. J. Psychol.*, **50**, 16–29.

THOULESS, R. H. (1931). Phenomenal regression to the "real" object, II. *Brit. J. Psychol.*, **22**, 1–30.

SIZE OF TEST PATCH AND SIMULTANEOUS CONTRAST

BY

PETER BURGH and G. C. GRINDLEY

From the Psychological Laboratory, University of Cambridge

Two experiments were done to find the effect on simultaneous brightness contrast of placing test patches in, as against outside, the fovea. When the display was magnified so that the test patches were outside the fovea no statistically significant effect was found. Very short exposures, to eliminate eye movements, did not appear to reduce contrast.

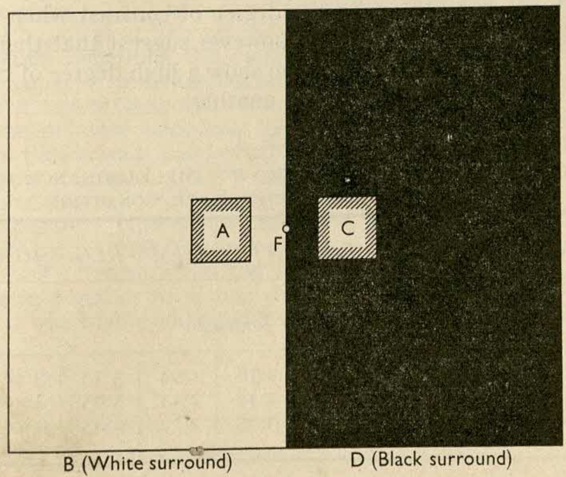
INTRODUCTION

It has usually been held that the familiar phenomenon of simultaneous contrast is partly central and partly peripheral (i.e. in this context, retinal) in origin. Reviews of the earlier literature are readily available (e.g. in Graham, 1934; or Woodworth, 1938). A recent review has been made by Diamond (1960).

If we suppose that the retinal component is important, it might be expected that the amount of contrast would be different inside and outside the foveal area, since there are so many well known structural and functional differences between the fovea and the periphery. Recent electro-physiological work on lateral inhibition in the retina (Barlow, FitzHugh and Kuffler, 1957; Wiesel, 1960; Barlow, 1953; Kuffler, 1953) suggests that there are changes in the size of the receptive field (both excitatory and inhibitory parts) and in the amount of inhibition, as we go from the centre towards the periphery. Other experiments (Hubel and Wiesel, 1959, 1960*a*, 1960*b*) have demonstrated a somewhat similar organization at higher levels in the central nervous system, and suggest a highly complex interaction of retinal and central factors. These experiments do not yet, however, give a clear indication of the kind of change to be expected in simultaneous contrast with variation of retinal area in the human being.

In this paper we are discussing the effect of magnifying the contrast-inducing display, so as to throw the test patches of Figure 1 wholly within or wholly outside the fovea, as opposed to viewing peripherally a display of constant size. The effect of very short exposures, so as to preclude eye-movements, is also considered.

FIGURE 1



Display and apparatus

In both the experiments to be described, the display to the subject was of the conventional type illustrated in Figure 1. Area (A) is a "grey" square seen in a "white" surround (B), while (C) is a "grey" square which is seen against a "black" surround (D). This display was produced by taking a piece of white cardboard and covering one half (D) with black paper. The squares (A) and (C) were windows cut in this cardboard, behind which there were other cardboard surfaces (A) and (C) whose luminance could be varied independently. The luminances of all surfaces were measured (from the direction of the subject's eye) with an Ilford S.E.I. Photometer. The windows were separated by an arbitrary distance equal to the length of their own sides, and the fixation mark (F) was symmetrically between them.

Details of the dimensions, conditions of viewing, and the luminances employed will be given in the accounts of the two experiments.

EXPERIMENT I

In the first experiment the display card measured 20 in. × 25 in., and the windows had sides of 1 in. and were separated by 1 in. The subject, using a head rest, viewed the card from a distance of about 6 ft. under three different conditions of magnification which altered the size of the squares (Fig. 1), in terms of visual angle, as follows:—

(*a*) through a prismatic monocular, giving squares of 5° visual angle;
(*b*) without optical apparatus (squares 1° visual angle);
(*c*) through a black tube at the far end of which there was a concave lens (squares 20′ visual angle).

Thus, on fixating the mark (F) both squares would throw their images outside the fovea in one extreme case, and onto it in the other.

The illumination of the three surfaces, (A), (B) and (C) was provided by three separate 150-watt lanterns with adjustable crossed polaroids. Those illuminating (A) and (B) were pre-set by the experimenter, while the control handle of that illuminating (C) was within reach of the subject's right hand while he was looking at the display.

The luminance of (A) was set at 2·5 ft.-L. and that of (B) at 10 ft.-L. Stray light caused the luminance of the "black" surround (D) to be about 0·2 ft.-L.

There were 10 subjects. Their instructions were to adjust area (C) until it "looked the same brightness" as area (A). Eight readings for each condition of magnification were given by each subject. Usual precautions were taken about ascending or descending values of the variable, and to show the three magnifications in different orders to different subjects. The complete session lasted about half an hour.

Results

The mean results are given in Table I. They are expressed in "degree of contrast," i.e. the ratio of the luminance of (A) to that of (C), at the point of subjective equality.

It is clear from inspection that these results do not show any general or statistically significant trend towards a change in the degree of contrast when the test patches are inside or outside the fovea. They do however suggest that there are differences between different subjects in that those who show a high degree of contrast with one size of test patch also tend to show it with another.

TABLE I

RATIO OF THE LUMINANCE OF THE STANDARD TO THE LUMINANCE OF THE VARIABLE FOR EACH SUBJECT UNDER EACH CONDITION

Subject	*KJB*	*JMO*	*GSJ*	*AOG*	*PAR*	*AJW*	*DJM*	*RLG*	*JGR*	*BM*	
Size of square	*Degree of contrast*										*Average*
5°	1·63	1·79	2·17	2·27	2·50	2·78	2·94	3·13	3·48	4·17	2·67
1°	1·15	2·00	2·09	3·34	2·39	3·48	2·23	3·05	4·30	3·13	2·72
20′	1·00	1·28	2·00	3·48	2·27	3·05	2·17	2·45	4·10	3·01	2·48

There are two main criticisms of this experiment. The first is that an alteration in the adjustment of the polaroids used caused a colour change as well as a brightness change; and the second is that even the most experienced observers doubted their ability to maintain fixation well enough for reliable observations on the 20′ squares. Small involuntary eye-movements could produce a kind of integration of after-images in each square, which could be mistaken for simultaneous contrast. Experiment II was designed to overcome both of these objections.

EXPERIMENT II

In some informal preliminary trials using a display in which both the test patches (A) and (C) were of equal reflectance, a large number of subjects (about 50) were asked to compare the contrast effect in continuous viewing by tungsten light with the contrast effect when the display was illuminated by an electronic flash of about one millisec. duration. A big majority reported that the contrast effect was greater with the flash. But, as the conditions of viewing were not standardized, a more controlled experiment was considered necessary.

Apparatus

The display was of the type described in Experiment I. The square windows had sides of 2½ in. and were separated by 2½ in. Behind the screen was a grey paper (C), which served as a standard, and a range of eight different grey papers (A), which could be displayed one at a time.

The display was illuminated by a 2 in. × 2 in. projection lantern, fitted with a 150-watt bulb and enclosed in a light-tight box, at a distance of 6 ft. from the display, and aligned and focused so that no direct light fell on the black half of the display screen.

The lantern was fitted with a Compur photographic shutter, which could be set to deliver flashes of 1/100 sec., or exposures of longer duration, by timing with a stop watch. A fixation point (F) was provided by making a hole in the display screen, directly between the test patches. Behind the hole was enclosed a small torch bulb.

The luminance of the surrounds and the test patches were measured by means of an Ilford S.E.I. Photometer. It was also ascertained with this photometer that any gradient of illumination over the white surround and over the test patches was negligible.

The luminance of the white surround (B) was 3·2 ft.-L., that of the black surround (D) was unmeasurable on the photometer, and was therefore less than 0·01 ft.-L. The luminances of the variable (A) lay between 0·80 and 4·35 ft.-L., while the standard (C) was 0·58 ft.-L. The experiment was done in a large room which was blacked out. Between readings, only the fixation point (F) was illuminated.

Method

The experiment was designed as a group experiment. Ten subjects were used; five in each of two sessions. They were drawn from the staff and research workers of the Psychological or Physiological Laboratories in Cambridge. The subjects were instructed to make judgements concerning the variable on a five-point scale as follows:—much brighter; brighter; equal or "different, but impossible to say in which direction the difference lies"; darker; much darker (than the standard).

Half the judgements were made from the front of the room, at a distance of 7 ft. from the display, where the stimulus patches subtended about 2° of visual angle, and half from the back of the room (distance 25 ft.), where the patches subtended about 30′. Likewise half the judgements were made on a 10-sec. exposure of the display, and half were made on a 1/100 sec. exposure. Thus each variable was presented once under each condition: 2° flash, 2° continuous, 30′ flash and 30′ continuous.

Before presenting each variable, the experimenter called out a letter, which the subjects wrote down together with their judgement; three letters were allotted to each variable, so that it was highly improbable that the subjects could identify any variable by its letter.

Results

The results given in Table II indicate the number of subjects showing definite evidence of contrast for each variable under each condition. It is evident from this table that under all conditions the majority of subjects showed a degree of contrast (the ratio of the luminances of (A) to (C)) of between 2:1 and 4:1.

TABLE II

NUMBER OF SUBJECTS (N = 10) REPORTING "A" TO BE "DARKER" OR "MUCH DARKER" THAN "C" UNDER EACH CONDITION.
(None of the subjects knew that "A" was, in fact, always lighter than "C.")

Size of square	*Exposure time*	*Ratio of luminances (A/C)*							
		1·2	2·3	3·0	3·1	3·3	3·7	4·3	7·2
2°	10 sec.	9	7	5	5	6	3	2	0
	1/100 sec.	8	6	3	5	3	6	3	0
30′	10 sec.	10	9	6	6	5	4	2	0
	1/100 sec.	9	8	7	5	5	4	3	0

The data are insufficient for detailed statistical analysis, but it is clear from inspection that these variations in the area of the test patches or in the duration of exposure do not greatly alter the order of magnitude of the contrast effect. Exposure of the display for a period short enough to preclude eye-movements does not eliminate contrast.

SUBJECTIVE COMMENTS

During these experiments we have been very much struck by a point which seems to have had little emphasis in the literature, namely the subjective difficulty of making a brightness match when the display is of this kind. Some comments by very experienced observers (all of them authors of scientific papers on various aspects of the psychology or physiology of vision) were:

"I can see that these two patches are the same colour, but I have the same *sort* of difficulty as in heterochromatic photometry" (H.B.B.).

"It depends how naive you want me to be; do you want me to say when I believe those patches are really of the same luminance, or to make some other sort of judgement?" (W.A.H.R.).

"Can't match them, because that on the black (surround) has a self-luminous quality" (R.L.G.).

"One begins to realize how complex the relations between contrast and constancy are" (O.L.Z.).

Less experienced subjects were, of course, less explicit about the nature of the difficulty; one, however, said:

"I attempt to adjust it so that the surfaces behind the two windows look as though they are parts of the same object" (K.P.W.). Apparent colour differences were reported by several subjects, as was a tendency to compare the brightness of the test patch, not with the comparison patch, but with the surround. There were very few who thought the judgements were at all easy.

In the second experiment, besides comments similar to those noted above, two of the subjects reported that under the 30′ condition the judgements were easier to make, or brightness differences were "more pronounced."

DISCUSSION AND CONCLUSION

The results of these experiments are mainly negative: no striking difference is found in degree of contrast for foveal and extra-foveal test patches.

The finding that there was little change in the average degree of contrast when the display was magnified may not, however, be a reliable indication of similarity of function in the retinal areas concerned. Leibowitz, Mote and Thurlow (1953) and Newson (1958) have shown the importance of close proximity between the test and inducing fields, and it might well be that the effect would be greater if the test patches were viewed peripherally, but kept at a constant size: some experiments by one of us (P.B. unpublished) suggest that under some conditions this is actually the case, and also that an increase of contrast is found towards the periphery.

Similarly, the fact that in Experiment II, when magnifying the test patches, the surround was also magnified, should have little bearing on the results. According to the authors quoted above, the difference in the size of the surrounds under the two conditions would have a negligible effect.

The rather large scatter between subjects, as illustrated in Table I, seems likely to be related to the subjective difficulty of the observations.

It may, in fact, be suggested that the subjects were being instructed to use criteria inappropriate to the situation; that they were being asked to make a judgement which, at least for some values of the variable, was not possible in the circumstances. The differences between subjects would suggest that they might have been using, perhaps unwittingly, different criteria for their judgements, and this is supported by comments such as those listed above.

The suggestion from Table I of large individual differences between subjects has been further investigated by one of us (P.B. unpublished), and the results tend to confirm these findings.

In view of these difficulties what is obviously wanted is a better method of measuring contrast.

We wish to express our thanks to Professor O. L. Zangwill, Dr. H. B. Barlow and Dr. F. W. Campbell for advice and discussion, and to the Medical Research Council, from whom one of the authors (P.B.) is in receipt of a grant.

REFERENCES

BARLOW, H. B. (1953). Summation and inhibition in the frog's retina. *J. Physiol.*, **119**, 69–88.

BARLOW, H. B., FITZHUGH, R., and KUFFLER, S. W. (1957). Change of organization in receptive fields of the cat's retina during dark adaptation. *J. Physiol.*, **137**, 338–54.

DIAMOND, A. L. (1960). A theory of depression and enhancement in the brightness response. *Psychol. Rev.*, **67**, 168–99.

GRAHAM, C. H. (1934). Vision, some neural correlates. In Murchison, C. (ed.). *A Handbook of General Experimental Psychology*, 829–77.

HUBEL, D. H., and WIESEL, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.*, **148**, 574–91.

HUBEL, D. H., and WIESEL, T. N. (1960*a*). Receptive fields of optic nerves in the spider monkey. *J. Physiol.*, **154**, 572–80.

HUBEL, D. H., and WIESEL, T. N. (1960*b*). Integrative action in the cat's lateral geniculate body. *J. Physiol.*, **155**, 385–98.

KUFFLER, S. W. (1953). Discharge patterns and functional organization of the mammalian retina. *J. Neurophysiol.*, **16**, 37–68.

LEIBOWITZ, H., MOTE, F. A., and THURLOW, W. R. (1953). Simultaneous contrast as a function of separation between test and inducing fields. *J. exp. Psychol.*, **46**, 453–6.

NEWSON, L. J. (1958). Some principles governing changes in the apparent lightness of test surfaces isolated from their normal backgrounds. *Quart. J. exp. Psychol.*, **10**, 82–95.

WIESEL, T. N. (1960). Receptive fields of ganglion cells in the cat's retina. *J. Physiol.*, **153**, 583–94.

WOODWORTH, R. S. (1938). *Experimental Psychology*. New York. Chapter XXII.

IMMEDIATE MEMORY IN DIALLING PERFORMANCE WITH AND WITHOUT SIMPLE REHEARSAL

BY

ALASTAIR HERON

From the Medical Research Council's Unit for Research on Occupational Aspects of Ageing, University of Liverpool

This experiment was designed to explore the effects of simple rehearsal on immediate memory in telephone dialling performance in a sample of 36 men aged 45 to 60 for whom data on intelligence, vocabulary, emotional stability and occupation were available.

The eight-digit messages were the same as those previously used by Conrad in a study of telephone operators which compared performance on dialling and keysender equipment. Using a Latin square design, subjects were required to dial 10 eight-digit messages in each of six conditions (three with auditory and three with visual presentation). Dialling was carried out (1) immediately the message had been presented; (2) after it had been once vocally rehearsed by the subject; (3) after it had been written down while being vocally rehearsed. A complete record was obtained of all performances, at rehearsal and at dialling.

It was found that with auditory serial presentation, immediate vocal rehearsal and immediate dialling are equally good and both superior to dialling which followed vocal rehearsal. With visual simultaneous presentation, immediate vocal rehearsal is clearly superior to dialling with or without rehearsal, which do not differ. Performance after written-down rehearsal was the worst in both modes of presentation.

A non-linear relationship between intelligence and performance was found, and acute individual differences in performance were observed among the subjects of the highest intelligence.

INTRODUCTION

In a recent paper on experimental psychology in the field of telecommunications, Conrad (1960*a*) refers briefly to the decay theory of immediate memory as "relevant to the problem of remembering a series of digits—say eight or nine—for long enough to dial them by telephone." He goes on to point out that decay theory is interesting in this context "because the telephone dial forces on the person dialling a series of delays between digits—of an average duration of nearly a second." The result of these enforced delays is, consistent with decay theory, that the chances of dialling the wrong number are increased (Conrad and Hille, 1958).

The experimental subjects in the studies reported by Conrad, in the paper quoted and elsewhere, are without exception young, with ages in the range 18 to 30 years. Although there does not appear to be any fall-off with age on the Digit Span subtest of the Wechsler Adult Intelligence Scale (Birren, personal communication), it does not follow that dialling performance will also remain constant: the effects of the enforced delay referred to by Conrad may be greater among older subjects.

The objects of the present experiment were as follows:—

(1) To ascertain whether simple rehearsal of the message was capable, despite the added delay involved, of improving performance at dialling as compared with immediate dialling.

(2) To obtain data from older subjects which could be compared with those of Conrad on young subjects.

(3) To explore, in a preliminary way, the differences in dialling performance, with and without simple rehearsal, which result from delivery of the message to the subject by (*a*) serial auditory presentation and (*b*) simultaneous visual presentation. (The first mode of presentation is of course identical with that involved in using the "Directory Enquiry" service; the second, as will be seen, was roughly analogous to looking-up the number in a directory oneself.)

METHOD

Apparatus

The subject was seated at a table on which was placed a conventional British G.P.O. dial telephone, the handset of which was not used. Under conditions of auditory presentation, messages were fed into a pair of headsets, one worn by the subject and one by the experimenter. The sequence of digits dialled was recorded electrically on a visual display placed in front of the experimenter. Vocal rehearsals by the subject were tape-recorded: a record of written rehearsals was obtained by using a modified cash-till, of the type still occasionally seen in village shops. Each amount is written down by the shop-keeper on a paper strip, through a small slit in the top of the till. The action of opening the cash drawer moves the paper on by means of a ratchet mechanism. By using a stripped-down version of this simple device it was possible to obtain a permanent record of what the subject wrote down when rehearsing, while still being able to remove what he had written before dialling commenced. (A possible disadvantage of this method, to which the attention of the writer has been drawn, is that the noise of operation may have impaired the performance of some subjects.)

The visual display was by means of an endless-loop film-strip, using automatic projection on a ground-glass screen. The eight-digit message was visible to the subject through a 1½ in. aperture, 9 in. in length, at a distance of 8 ft., 45 degrees to the left of the sitting position.

Stimulus material

Six lists of 10 eight-digit messages were prepared from the supply of 80 such messages employed by Conrad (1958). Each digit occurred an equal number of times in each serial position, obviously easy phrases being avoided. For auditory presentation, these lists of messages were tape-recorded at a rate of 72 digits per minute; the visual display of each message was of 8 sec. duration. A standard interval of 40 sec. between messages was used under both conditions of presentation.

TABLE I

DETAILS OF EXPERIMENTAL POPULATION

	Age	*P.M.*	*MHV*	*H–1*	*H–2*
Mean	51·4	43·1	34·3	7·1	5·8
Range	45–60	28–56	27–41	0–13	0–12
σ	3·1	6·8	4·2	3·2	2·8

P.M.—Progressive Matrices 1938; MHV—Mill Hill Vocabulary, Senior Synonyms; H–1, H–2—Heron Two-part Personality Inventory, scores for emotional instability and unsociability respectively.

Subjects

The subjects were 36 male members of the Medical Research Council voluntary subject panel at Liverpool. As will be seen in Table I, the mean age is 51·4 years; the group mean on Progressive Matrices 1938 is equal to that given by Raven for the 50th percentile up to the age of 30, and it falls between the 75th and 90th percentiles for age 50. The mean score on vocabulary is equal to that for the 75th percentile at all ages between 25 and 55, while that on Part 1 of the Heron Inventory (emotional maladjustment) lies at the upper end of the "probably well-adjusted" range (0–7). On Part 2 the mean score is at the lower end of the "unsociable" range (6 and up).

Design

There were six experimental conditions, as follows:

Auditory presentation		*Visual presentation*
1	Immediate dialling	4
2	Vocal rehearsal before dialling	5
3	Simultaneous vocal and written rehearsal before dialling	6

These six conditions were randomized in six different 6 × 6 Latin squares, and each subject was tested under each condition with a different list of messages.

Procedure

In conditions 1, 2 and 3, the subjects were told first "You will hear the operator say 'The number is . . .' and then eight figures"; similarly in conditions 4, 5 and 6 they were told first "I will say to you 'The number is . . .' and then you will see the eight-figure number here (*pointing at display*) for a few seconds."

In conditions 2 and 5, the subjects were then told "(Immediately after the last figure) (Immediately the number disappears) but NOT before, you are to say out loud what you have (heard) (seen), and then to dial the whole number. DO NOT DIAL UNTIL YOU HAVE FINISHED SAYING IT OUT LOUD."

In conditions 3 and 6, the subjects were told "(Immediately after the last figure) (Immediately the number disappears) but NOT before, you are to *write down* the whole number *saying it out loud as you do so*. You will *not* be able to copy what you have written. DO NOT DIAL until you have finished writing and saying the whole number."

At this point each subject was given three practice trials under the appropriate display and response conditions.

Under all conditions, the remaining instructions were as follows: "When you are dialling, please do not stop or hesitate, but there is no need to hurry—just work at the pace of the dial mechanism. (PAUSE) If you are doubtful or think you have forgotten one of the figures, just dial whatever you think is the most likely one. DO NOT STOP the smooth operation of your dialling. (PAUSE) Have you any questions? (PAUSE) Right—let's begin."

(*Note.*—These final general instructions were not repeated beyond the third of the six conditions for any given subject, since such repetition could easily be annoying.)

Administration

Two female members of the M.R.C. staff shared the conduct of this experiment equally between them. Twelve subjects were tested during afternoons, 24 during evenings. Possible effects of these two extraneous variables (experimenter and time of day) were tested, by analysis of variance in each case, using total number of correctly dialled messages under all six experimental conditions; no significant differences were found. It should also be noted that since all these subjects participated on the same 2-hr. visit in another experiment, care was taken to ensure that half the subjects took part in each experiment first before going to the other.

RESULTS

A message was scored as correctly reproduced only when all eight digits were given in the correct order. No significant difference was found between the six Latin squares when the total number of messages dialled correct was used as a test score in an analysis of variance and they were therefore combined in all examinations of condition effects.

Comparison within each display condition of results under different response conditions

In Table II will be seen the results obtained under conditions of auditory and of visual presentation.

TABLE II

MEAN NUMBER OF MESSAGES CORRECTLY DIALLED OR CORRECTLY RECALLED ON REHEARSAL (out of a possible 10)

	Serial auditory presentation	*Simultaneous visual presentation*
Vocal rehearsal	3·81	6·08*
Immediate dialling	4·03	4·81
Dialling after vocal rehearsal . .	3·19*	4·92

* 3·19 is smaller ($p < 0·01$) and 6·08 larger ($p < 0·001$) than the other two figures in their respective columns.

The first point to dispose of is the uniformly better performances under conditions of visual simultaneous presentation of the eight digits. Since there are no grounds for assuming any equivalence of stimulus, no detailed cross-comparisons are justified. What can be compared, however, are the two patterns of within-condition differences. With auditory serial presentation, immediate vocal rehearsal or immediate dialling without rehearsal are equally good, whereas with visual simultaneous presentation, immediate vocal rehearsal is significantly superior to immediate dialling. With auditory presentation, immediate rehearsal and immediate dialling are both superior to dialling which followed vocal rehearsal. With visual presentation, this is not the case.

To avoid confusion, the results obtained when dialling after combined vocal and written-down rehearsal have been omitted from the table. In both conditions of stimulus presentation, these are the poorest dialling performances (2·58 and 3·92 respectively) possibly partly as a result of noise distraction.

Effects on dialling of spontaneous grouping during vocal rehearsal

As might be expected, many subjects grouped the digits spontaneously during vocal rehearsal; of those who did, the majority broke the message into two 4-digit groups. No other split was used by a sufficiently large number of subjects to justify separate treatment in statistical analysis of the data. It will be seen in Table III

TABLE III

EFFECTS OF SPONTANEOUS GROUPING OF DIGITS ON DIALLING AFTER VOCAL REHEARSAL

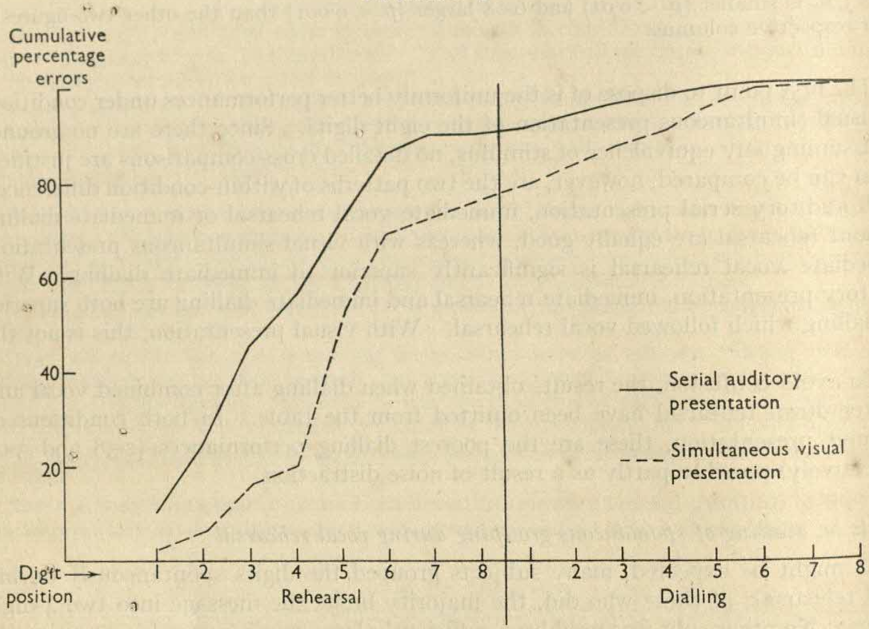
	Serial auditory presentation				*Simultaneous visual presentation*			
	Messages correctly dialled	*Per cent.*	*Not correctly dialled*	*Per cent.*	*Messages correctly dialled*	*Per cent.*	*Not correctly dialled*	*Per cent.*
Grouping in fours . .	60	(38·0)	98	(62·0)	109	(61·6)	68	(38·4)
Others . .	55	(27·2)	147	(72·8)	65	(35·5)	118	(64·5)
	115		245		174		186	

that dialling following simple vocal rehearsal was significantly better among those dividing the message into four-digit groups; this was the case with both auditory and visual presentation, but the effect was much more marked in the latter, the success ratio actually reversing for grouped performance.

The location of initial error

Since a complete record was obtained of all performances by all subjects under every condition, it has been possible to examine the serial development of the patterns of initial error. As will be seen in Figure 1, the two modes of presentation result in somewhat different cumulative error patterns.

FIGURE 1



Vocal rehearsal followed by dialling: cumulative percentage of errors at each point in the digit sequence.

With serial auditory presentation of the messages, errors occur in a smooth progression through digit positions 1 to 6, with only a slight spurt at position 3. Ninety per cent. of the eventual total number of errors have been made during rehearsal before dialling commenced. This pattern of error contrasts sharply with that shown for simultaneous visual presentation, where less than 20 per cent. of errors occur in the first four positions, half occur at position 5 and 6, and only 76 per cent. have been made before dialling commences.

Of the 136 and 217 messages correctly vocally rehearsed under auditory and visual conditions respectively, 111 and 173 (81·6 and 79·7 per cent.) were then correctly dialled: the further loss during dialling was in both cases accounted for by initial errors distributed over the first five digits.

Of the 118 and 175 messages which were correctly rehearsed both vocally and in writing under auditory and visual conditions respectively, 92 and 142 (78·0 and 81·1 per cent.) were then correctly dialled.

Errors

The most clearly-definable type of error is that of transposition, involving two or three adjacent digits. This type accounted consistently for just over a quarter of the errors made when dialling under auditory presentation conditions. With visual presentation, the same proportion obtained for dialling after rehearsal, but in the case of immediate dialling, transposition errors were only 16 per cent. of the total. In immediate vocal rehearsal, the proportion was about a third for both modes of presentation.

TABLE IV

Transposition Errors

	Serial auditory presentation	*Simultaneous visual presentation*
	Percentage of total errors	*Percentage of total errors*
Immediate dialling	27·9	16·0*
Vocal rehearsal	32·7	35·5
Dialling after rehearsal ..	26·1	29·0

* Significantly different ($p < 0{\cdot}01$) from the other figures in the column.

Intelligence

Inspection of bi-variate plots of the association between non-verbal intelligence and performance indicated considerable non-linearity. The 36 subjects were therefore divided into three groups: Highs (48–56), N = 13; Mediums (40–47), N = 10; and Lows (28–39), N = 13. (The unequal division was dictated by tied scores around the upper and lower one-third cutting points). The resulting analysis is set out in Table V, from which it will be seen that the association with non-verbal intelligence varies with the mode of presentation and also with the performance under consideration. Since it can be predicted that subjects of high intelligence should perform better than those of lower intelligence on a task involving memory for digits of a considerable span, a one-tailed test has been used in assessing significance.

TABLE V

Association between Intelligence and Performance at Rehearsal and Dialling (Messages Correct)

	Serial auditory presentation			*Simultaneous visual presentation*		
	H	*M*	*L*	*H*	*M*	*L*
Vocal rehearsal	5·38	3·40	2·54	7·38	5·20	5·46
Immediate dialling	6·38	3·30	2·23	6·69	2·70	4·54
Dialling after rehearsal ..	4·69	2·70	2·08	6·61	4·20	3·77

Differences between performances (rows) tested by Wilcoxon matched-pair signed-ranks test; those between intelligence levels (columns) by Whitfield's tau method. For results see text.

Immediate dialling. Highs are superior to Mediums and Lows for both modes of presentation ($p < 0{\cdot}01$).

Immediate rehearsal. Highs are superior to Lows for both modes of presentation ($p < 0{\cdot}03$) and also to Mediums for visual only ($p < 0{\cdot}05$).

Dialling after rehearsal. As for immediate rehearsal.

With one borderline exception (I.D. visual) Mediums and Lows do not differ significantly between themselves.

Comparison of performance within each level of intelligence. If the table is read downwards, it is possible to see whether the various performances differ among themselves within every level of intelligence. In fact, they do not. Since one might reasonably predict that immediate vocal rehearsal would give results superior to those on either immediate dialling or on rehearsed dialling, a one-tailed test can be used for these two comparisons.

In the case of auditory presentation, the hypothesis is not sustained that performance on immediate rehearsal should be superior to that on immediate dialling, at any level of intelligence; for visual presentation, it is sustained ($p < 0{\cdot}03$) at medium and low levels of intelligence but not at the highest level.

Performance at immediate rehearsal is significantly better ($p < 0{\cdot}03$) than at dialling after rehearsal, at every level of intelligence in both modes of presentation except Mediums, auditory—and the difference here is in the consistent direction ($p = 0{\cdot}07$).

When comparing immediate dialling with dialling after rehearsal, a two-tailed test was used in the absence of any sound basis for a directional hypothesis. Consistent with the results already presented for the whole group of subjects ($N = 36$), no significant differences were found at any level of intelligence with visual presentation. With the auditory mode, there was a significant difference for the Highs ($p < 0{\cdot}03$); those for Mediums and Lows were in the consistent direction but not significant at the 5 per cent. level. From this it seems that the observed difference for the whole group is largely attributable to those subjects of the highest intelligence.

Characteristics of best and worst performers equated for non-verbal intelligence

Since the 22 subjects with the highest intelligence included eight of the best and also eight of the worst performers at immediate dialling under conditions of serial auditory presentation, it seemed worthwhile to examine other characteristics of these two groups. Table VI sets out the available information: it may be noted that the groups are equated for chronological age as well as for non-verbal intelligence, and that the two groups are consistently good and poor performers respectively in all three performances under both modes of message presentation.

TABLE VI

MEAN SCORES OF TWO GROUPS, EQUATED FOR AGE AND NON-VERBAL INTELLIGENCE

				Auditory			*Visual*			*Personality tests*				
	Age	*P.M.*	*Voc.*	*V.R.*	*I.D.*	*R.D.*	*V.R.*	*I.D.*	*R.D.*	H_1	H_2	W_1	W_2	W_3
Good performers	50·0	45·9	26·0	6·9	7·5	5·9	8·6	7·5	7·1	8·2	5·9	8·0	7·5	*2·5**
Poor performers	50·1	46·2	23·9	2·0	1·9	1·6	3·6	1·4	2·7	7·2	5·2	8·6	8·5	*3·7**

* Significantly different, $p = 0{\cdot}02/0{\cdot}01$.

Personality differences. The only significant difference between these two groups of good and poor performers is in their scores on Wesley 3: this is based on a group of items from the Wesley Personality Rigidity Inventory which have been found (Chown, 1961) to define a factor associated with "liking for established routine." The group of poor performers at dialling and recall appear to be more "rigid," in this defined use of the term.

Occupational differences. The possibility that good performance—when age, intelligence and personality factors have been dealt with—might depend to some extent on what a subject was used to, received but slight support from inspection of occupational differences.

DISCUSSION

Age differences in performance

Conrad (1960*b*) reports the results obtained from three experiments using naval ratings as subjects, the first two of which employed eight-digit messages. Auditory presentation was used, but at 120 digits per minute instead of the 72 d.p.m. used in the present experiment. No dialling was involved.

The average percentage of messages correctly recalled was about 60, as compared with the 40 per cent. of the older subjects at vocal rehearsal. This difference may be accounted for in part by the slower rate of presentation, since Conrad and Hille (1958) report 62 per cent. for a presentation rate of 90 d.p.m. and 43 per cent. for one of 30 d.p.m., again with young subjects but this time telephone operators. It seems unlikely, however, that the relatively small difference in presentation rate between 90 and 72 could wholly account for the present results, and an age-difference must at least be suspected.

In another study (Conrad, 1958) young telephone operators dialled correctly 50 per cent. of the eight-digit messages presented at 100 d.p.m. Again the difference between this and the 40 per cent. obtained from older subjects may be partly due to the slower rate of presentation.

The interaction between differences in age and in rate of presentation clearly requires separate investigation.

The effects of simple vocal rehearsal

It is quite clear that simple vocal rehearsal of an 8-digit message adversely affects the performance of older subjects as compared with immediate dialling, using serial auditory presentation of the message. This may be compared with the finding of Conrad (1960*b*) that a 5-sec. period for silent rehearsal did not affect the recall of his young naval ratings. But it has been pointed out that silent rehearsal can be faster than vocal, and the possibility must therefore be considered that two or three rapid silent rehearsals may have been enough to avoid decrement in the young subjects, while a single vocal rehearsal on the part of the older ones either failed to do so or actually made matters worse.

Some support for this is to be found in the results obtained using an 8-sec. display for simultaneous visual presentation. Clearly some subjects may have used this period for several rapid silent rehearsals, and this may partly explain the significantly better average performance at vocal rehearsal. This advantage does not fully extend to the subsequent dialling, but may have been enough to prevent it being no worse than was observed for immediate dialling without vocal rehearsal.

Taking the results as a whole, it seems reasonable to suggest that simple vocal rehearsal is contra-indicated: it does not help, and it may hinder.

Implications for decay theory

The results of the present experiment are generally in accord with decay theory. It is however surprising that under conditions of serial auditory presentation the delay enforced by dialling does *not* affect performance as compared with immediate rehearsal. This is, however, similar to the result obtained by Conrad (1958) when comparing performance on dialling with that on a keysender, which does not enforce a delay. Conrad found that the difference, although in the expected direction, was not significant, probably because of the much greater variation in dialling performance as compared with keysending.

On the other hand, as already noted, the difference between immediate recall and immediate dialling for young subjects reported by Conrad is that between the 60 per cent. correctly recalled by his naval ratings and the 50 per cent. correctly dialled by the telephone operators (one can here hazard the guess that the latter figure might have been slightly lower had the subjects again been naval ratings).

This leaves us with an age-difference in relation to decay theory which invites some sort of hypothesis by way of explanation. One possibility is that the older subjects have already forgotten so much before dialling commences that the additional delay imposed by the dial mechanism does not further affect their performance to any significant extent. This received some support from the fact that 90 per cent. of the errors made in rehearsed dialling had occurred during vocal rehearsal before dialling commenced. The writer has not so far been able to establish whether a smaller proportion of errors at the recall stage has been found with young subjects.

Kay (1959) and Welford (1958) have drawn attention to the apparently greater effect on short term retention by older subjects of any form of intervening activity. What one is now asking is whether the arrival of another digit in a series constitutes "interference" in older subjects so great and so cumulative that in a series of eight digits the effect is enough to make the dialling delay relatively unimportant. It has been argued that this arrival of later digits in a series has its effect simply by delaying recall, and therefore that decay is wholly adequate as an explanation for observed decrement. This argument does not seem wholly acceptable and further experimental work is needed to clear the point. Here we must observe that with simultaneous visual presentation the enforced delay due to dialling did have a marked effect; as already noted, considerable silent rehearsal was possible during the display, and since it was simultaneous there is less question of "interference"; delay then appears as the major factor in the situation.

There seems to be a possibility that when age-differences are taken into account, performance involving very short-term memory may require both "interference" and "decay" theories to explain it fully.

Individual differences

The choice of eight digits as the length of message made it likely that differences in intelligence would affect performance on recall and probably also on dialling. As we have seen, however, eight of the poorest performers were found among those subjects of the highest intelligence where eight of the best performers were also found. Whatever may be the nature of the association between immediate memory for digits among young subjects, it is clear that by the age of 50 years one must also reckon with other determining factors in performance. In this connection while bearing in mind the potential relevance of occupational experience it is clearly unwise to jump to conclusions.

The author wishes to express his appreciation to Mr. D. E. Broadbent and Dr. R. Conrad for their sustained interest in this experiment: discussion with them both has been very valuable. He is also grateful for the help of Mrs. Brigid Brechling and Miss Margaret McEwan, who conducted the experiment; the latter was also largely responsible for the computational work which resulted.

REFERENCES

CHOWN, S. M. (1960). A factor analysis of the Wesley Rigidity Inventory: its relationship to age and nonverbal intelligence. *J. abnorm. soc. Psychol.*, **61**, 491–4.

CONRAD, R. (1958). Accuracy of recall using keyset and telephone dial, and the effect of a prefix digit. *J. appl. Psychol.*, **42**, 285–8.

CONRAD, R. (1960*a*). Experimental psychology in the field of telecommunications. *Ergonomics*, **3**, 289–95.

CONRAD, R. (1960*b*). Very brief delay of immediate recall. *Quart. J. exp. Psychol.*, **12**, 45–7.

CONRAD, R., and HILLE, B. A. (1958). The decay theory of immediate memory and paced recall. *Canad. J. Psychol.*, **12**, 1–6.

KAY, H. (1959). Theories of learning and aging. Chapter 18 in Birren, J. E. (Ed.), *Handbook of Aging and the Individual: Psychological and Biological Aspects.* Chicago.

WELFORD, A. T. (1958). *Ageing and Human Skill.* London.

SHORTER ARTICLES AND NOTES

AN INVESTIGATION OF CHANGES IN AUDITORY SENSITIVITY DURING THE PERFORMANCE OF A MENTAL TASK

BY

J. R. SYMONS and C. K. MACKAY

From the Department of Psychology, University of Aberdeen

Auditory thresholds were measured while the subject was relaxed and while he was doing mental arithmetic. Thresholds were lowered under the second condition. Some possible explanations are briefly considered.

INTRODUCTION

In a well known paper, Bruner (1957) suggests that some reinforcement might be essential for a stimulus input to reach a level of activation where a "fit" could be made between the input and a previously established model. Although this idea would be technically difficult to test, he believes that what needs to be shown is that there is an onset of arousal impulses from the reticular formation while the observer is trying to figure out the nature of an input and a diminution of this activity following the identification of the object. In other words he would like to demonstrate that Bartlett's "effort after meaning" is dependent for its success on reinforcing stimulation from the reticular system.

The problem posed by Bruner brings to mind an early experiment by Travis (1926) who measured the auditory thresholds of his subjects before, during, and after problem solving, serious reading, memorizing poetry, and reading a love story. Of his 35 subjects 27 recorded lower thresholds during the mental activity. It is possible that the small increase in auditory sensitivity indexed a general arousal of the organism, which is what Bruner would predict. In any event it was thought worth while to carry out the following small study.

METHOD

Apparatus

A pure tone oscillator provided a signal of 1,000 c.p.s. The attenuator scale ranged from 0 dB to minus 80 dB. The output power at the earphones at 0 dB was 1 microwatt.

Each subject was given a printed sheet of 15 questions with spaces provided for answers. The questions ranged in difficulty from question 1 "If you can buy one apple for 4d. how much would you pay for three apples," to question 15: "Each letter represents a number between 0 and 9 inclusive. No two letters represent the same number, and in the case of the articles bought no two numbers represent the same letter. Solve the puzzle given that E equals 5.

	£	s.	d.
PP 8317 at B/R d.	P	S	R
AA 1547 at A d.		M	A
PTV 25347 at P d.		V	P
	B	N	PT."

Subjects

Thirty subjects were randomly divided into an experimental group of 20 and a control group of 10. Ages ranged from 8 to 62 years. The mean age of the experimental group was 27·8 years with a standard deviation of 15·7, and that of the control group 26·5 years with a standard deviation of 13·6. There were seven males and 13 females in the experimental group, and five males and five females in the control group.

Procedure

The light adapted subject, who was seated comfortably at a table, was told that the experiment was concerned with hearing ability. In the normal series (no mental task) he was instructed to look at a white sheet of paper on the table in front of him, and to concentrate on listening for the tone. In the stimulus series the subject was given a list of problems and was told to work through them, writing the answers down in the space provided. It was emphasized that of the two tasks the response to the onset of the tone was the most important, but that as many problems should be solved as possible. In both the normal and the stimulus series the subject responded to the tone by raising the index finger of his left hand.

The experimenter used a modified method of limits. The gain was set at randomly determined positions between 5 and 20 dB. below the threshold of the subject established in the practice session. After a verbal signal the gain was continuously increased at a constant rate of 2 dB. per sec. until the subject responded, whereupon the gain was immediately returned to the next starting position. Catch trials were introduced into both the practice and the experimental sessions. No responses to catch trials occurred during the experiment. The sequence of events can best be summarized as follows:—

Experimental group:

Practice period	10 min.
A. Normal series 1, 10 threshold estimates	10 min.
Rest..	2 min.
B. Stimulus series, 10 threshold estimates	10 min.
Rest..	2 min.
C. Normal series 2, 10 threshold estimates	10 min.

Control group:

The sequence of events was identical except that a normal series replaced the stimulus series.

RESULTS

The mean auditory thresholds obtained in each of the three series, normal series 1 and 2 and the stimulus series, for each of the 20 subjects of the experimental group are given in Table I, as are the standard deviations. Table I also includes the number of mental arithmetic problems correctly solved during the stimulus series. The corresponding data collected from the control group are shown in Table II. Since there is no stimulus series for this group there is no arithmetic score. Table III shows the means of these threshold values for each condition for both experimental and control groups. It also shows the differences between these means. These differences were tested, using Fisher's t test. The three differences obtained by the experimental group were significant at the 0·01 level. The differences obtained by the control group were not significantly different from zero. There was a negative correlation of 0·36 between the arithmetic score and the magnitude of the difference between the means of the first normal series and the stimulus series, with age "partialled out."

TABLE I

EXPERIMENTAL GROUP

Subject age and sex	*Means of* 10 *threshold determinations and S.D.'s of these determinations*						*Mental Arith. score*
	A. Normal 1		*B. Stimulus*		*C. Normal* 2		
	Mean	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*	
24, female	−51·8 dB.	1·3	−55·2 dB.	1·5	−53·7 dB.	1·2	7
24, female	−51·1 dB.	1·5	−52·7 dB.	1·9	−51·8 dB.	2·6	7
27, female	−55·3 dB.	0·8	−57·0 dB.	1·3	−54·8 dB.	1·1	12
16, female	−54·0 dB.	1·0	−55·2 dB.	2·7	−53·6 dB.	1·0	5
59, female	−45·4 dB.	1·5	−46·6 dB.	0·8	−45·4 dB.	0·8	10
31, female	−54·1 dB.	1·1	−57·7 dB.	1·1	−55·0 dB.	1·1	11
13, male	−53·2 dB.	1·5	−56·0 dB.	1·4	−52·9 dB.	0·9	6
13, female	−52·5 dB.	1·3	−54·7 dB.	1·0	−53·6 dB.	1·8	6
37, female	−52·6 dB.	0·8	−53·3 dB.	1·0	−53·0 dB.	0·8	8
16, male	−53·4 dB.	0·8	—56·3 dB.	1·1	−52·2 dB.	1·0	7
16, female	−52·4 dB.	0·8	−54·4 dB.	1·1	−52·9 dB.	0·5	7
16, female	−37·7 dB.	1·7	−41·5 dB.	1·1	−40·6 dB.	1·3	5
59, male	−41·0 dB.	0·6	−42·8 dB.	0·9	−41·9 dB.	0·7	5
20, male	−53·2 dB.	0·8	−53·8 dB.	1·2	−53·2 dB.	0·6	5
27, male	−52·3 dB.	1·0	−54·6 dB.	0·9	−53·4 dB.	0·8	4
62, male	−40·9 dB.	0·6	−44·6 dB.	0·6	−41·1 dB.	0·5	4
45, male	−41·9 dB.	1·7	−45·9 dB.	1·0	−42·7 dB.	1·4	4
16, female	−51·0 dB.	0·0	−53·9 dB.	0·8	−52·6 dB.	0·9	2
14, female	−52·3 dB.	0·9	−53·9 dB.	0·9	−52·6 dB.	0·9	4
25, male	−51·0 dB.	0·6	−52·8 dB.	0·9	−52·1 dB.	0·7	7

TABLE II

CONTROL GROUP

Subject age and sex	*Means of* 10 *threshold determinations and S.D.'s of these determinations*					
	A. Normal 1		*B. Normal* 2		*C. Normal* 3	
	Mean	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*
27, male	−53·2 dB.	0·8	−52·6 dB.	0·7	−52·6 dB.	1·6
27, female	−55·3 dB.	1·0	−55·2 dB.	0·7	−55·2 dB.	0·5
24, male	−51·6 dB.	1·1	−51·4 dB.	0·7	−51·2 dB.	0·8
26, male	−51·2 dB.	0·4	−51·1 dB.	0·5	−51·0 dB.	0·7
29 male	−48·3 dB.	1·8	−48·2 dB.	1·5	−48·4 dB.	1·7
27, female	−52·2 dB.	1·5	−51·8 dB.	1·8	−51·6 dB.	1·7
8, female	−44·1 dB.	1·0	−43·6 dB.	1·5	−44·3 dB.	1·0
27, male	−56·9 dB.	1·9	−57·4 dB.	2·3	−57·3 dB.	1·6
10, female	−51·5 dB.	1·0	−51·7 dB.	1·0	−51·3 dB.	0·6
62, female	−37·8 dB.	1·0	−37·5 dB.	0·6	−37·5 dB.	0·6

TABLE III

Experimental group			*Control group*		
Mean absolute auditory thresholds					
20 *subjects*			10 *subjects*		
A.	Normal series 1	−49·8 dB.	A.	Normal series 1	−50·2 dB.
B.	Stimulus series	−52·1 dB.	B.	Normal series 2	−50·1 dB.
C.	Normal series 2	−50·5 dB.	C.	Normal series 3	−50·4 dB.
Differences					
A—B.	2·3 dB.		A—B.	0·1 dB.	
B—C.	1·6 dB.		B—C.	0·3 dB.	
A—C.	0·7 dB.		A—C.	0·2 dB.	

DISCUSSION

A number of explanations of varying degrees of plausibility could be suggested to account for the small improvement in auditory sensitivity which accompanied the mental activity of the experimental group during the stimulus series. It is possible that there is a general increase in muscle tonus occurring under these conditions which effects the peripheral auditory mechanism of the middle ear. Or it could be argued that there was a tendency for the subject to react more rapidly during the stimulus series and thus shorten the time of exposure to the mental arithmetic problems. But since the experiment was suggested by Bruner's speculations on the function of the reticular activating system in "effort after meaning" another explanation should be considered. It is possible that any increase in muscle tonus accompanying this sort of mental activity would increase the activity level in the ascending brain stem facilitatory mechanisms. There is evidence that of all the sensory systems proprioception is particularly efficient in this respect. An explanation in terms of general arousal, an explanation which was advanced many years ago to explain the facilitatory effect of heteromodal stimulation, is congruent with Bruner's belief that some reinforcement is needed for a stimulus input to reach a level of activation where a "fit" can be made between input and "model."

But it will be noted that the difference between the means of the first and third series (normal 1 and normal 2) recorded by the experimental group is significant. While this result confirms that of an earlier experiment (Symons, 1949) in which the facilitatory effects of heteromodal stimulation on the auditory threshold was found to outlast the heteromodal stimulus by at least 7 min., it does not fit Bruner's requirement of a rapid fall in activation once the influx of activity has performed its function. The picture of the brain stem has become so vastly complicated since Moruzzi and Magoun (1949) published their pioneering paper on the function of the reticular formation that it would not be difficult to suggest how Bruner's requirement might be met. But this is not the place for such speculation.

The small negative correlation between the number of problems solved and the magnitude of the threshold shift suggests that the distribution of the subject's attention between the two tasks in the stimulus series is a critical factor. The work of Gregg and Brogden (1952) illustrates this point most clearly, and with their results in mind the subject was told that the auditory task was of more importance than the problem solving.

In conclusion, there is agreement between these experimental results and those reported by Travis. If the explanation which relies on the ascending facilitatory system of the brain stem is accepted the results of the two experiments give some support to Bruner's thesis. But physiologizing of this kind can all too easily leave solid ground, and for those thus tempted a re-reading of Brodal's Henderson Trust Lecture (1957) is a salutary exercise.

These data were first reported by the junior author in a thesis submitted for the Ed.B. degree, University of Aberdeen.

We gratefully acknowledge the technical assistance and advice given by Mr. C. J. S. Stuart, Senior Technician in the Psychology Department.

REFERENCES

BRODAL, A. (1957). *The Reticular Formation of the Brain Stem: Anatomical Aspects and Functional Correlation.* Edinburgh.

BRUNER, J. S. (1957). Neural mechanisms in perception. *Psychol. Rev.*, **64**, 340–58.

GREGG, L. W., and BROGDEN, W. J. (1952). The effect of simultaneous stimulation on absolute auditory sensitivity. *J. exp. Psychol.*, **43**, 179–86.

MORUZZI, G., and MAGOUN, H. (1949). Brain stem reticular formation and activation of the EEG. *Electroenceph. clin. Neurophysiol.*, **1**, 455–73.

SYMONS, J. R. (1949). Changes in auditory acuity through the simultaneous stimulation of other sense organs. *Unpublished B.A. thesis. Reading University.*

TRAVIS, L. E. (1926). Changes in auditory acuity during the performance of certain mental tasks. *Amer. J. Psychol.*, **37**, 139–42.

A FURTHER STUDY OF PARTIAL REINFORCEMENT IN THE TURTLE

BY

R. C. GONZALEZ and M. E. BITTERMAN
From the Department of Psychology, Bryn Mawr College

Two matched groups of mature painted turtles, *Chrysemys picta picta*, were trained in a simple runway, one with partial and the other with consistent reinforcement, following which both groups were extinguished. The partially reinforced animals ran more slowly in acquisition, but showed somewhat greater resistance to extinction. The results are compared with those obtained in analogous experiments with other animals.

INTRODUCTION

This is one of a series of comparative studies designed to explore the phyletic generality of certain theoretically significant phenomena of learning which have been established in work with the rat, hitherto the principal infrahuman subject of research on learning (Bitterman, 1960). One of these phenomena is the paradoxical effect of partial reinforcement on resistance to extinction. When work with the fish cast doubt upon the phyletic generality of the effect (Wodinsky and Bitterman, 1959, 1960; Longo and Bitterman, 1960; Gonzalez, Longo and Bitterman, 1961), it seemed desirable to examine some intermediate forms, and an experiment was performed with some very young turtles (*Pseudemys scripta elegans*) in which, again, the paradoxical effect failed to appear (Eskin and Bitterman, 1961). A similar experiment with some adult turtles is reported here.

METHOD

The subjects were 28 painted turtles (*Chrysemys picta picta*) selected from a larger group of about 50 animals which were collected in southern New Jersey during the summer months of 1960. They ranged in size (width) from 4 to 6 in. The animals spent the late summer and the early fall roaming freely in a small temperature-controlled greenhouse equipped with several shallow pools, an environment in which they seemed to thrive. During this time, there were exploratory experiments with 10 of the animals in a trial runway under varying conditions of deprivation and reward. In November, 39 animals with no previous training were moved to a temperature-controlled room in the main laboratory, where the present experiment was performed.

In the laboratory, the animals were placed on a 21-hr. water-deprivation schedule. They lived on dry paper-towelling in individual, slate-bottomed, 5-gal. aquariums. Once each day, about 2 hr. after the daily experimental session, each animal was transferred for 30 min. to another 5-gal. aquarium, containing shallow water, in which it received its daily ration of ground beef. (Water-deprivation and water-reward were used, because food did not constitute a sufficiently dependable incentive.) Artificial sunlight supplemented the normal room illumination for a period of 2 hr. each day.

The apparatus was a wooden runway, 8 in. wide, 10 in. high and 6 ft. long, with a U-turn at the far end leading to a large tank of water. From the end-compartment of the runway, the animal could slide directly into deep water, and from there, if it chose, it could find its way to a region of shallow water. Provisions were made for the insertion of a wooden barrier between the end-compartment and the tank, which prevented entry into (or sight of) the tank. About 4 ft. from the starting compartment of the runway, there was a sliding door which could be inserted behind the animal to prevent retracing. About 5 ft. from the starting compartment, at a point from which the animal could see neither the tank nor the barrier which prevented entrance into the tank on non-reinforced trials, a light-beam crossed the runway and fell on a photocell recessed in one wall.

Each animal was given a single trial in the runway each day. A Standard Electric timer was activated when the animal was placed in the starting compartment and stopped by a photocell-relay when the light-beam was interrupted. The experimenter then closed the retrace-door, recorded the running time, and reset the clock. On reinforced trials, the animal was returned to its aquarium 1 min. after it had entered the water. On non-reinforced trials, it was detained in the end-compartment between the retrace door and the barrier for an equal length of time.

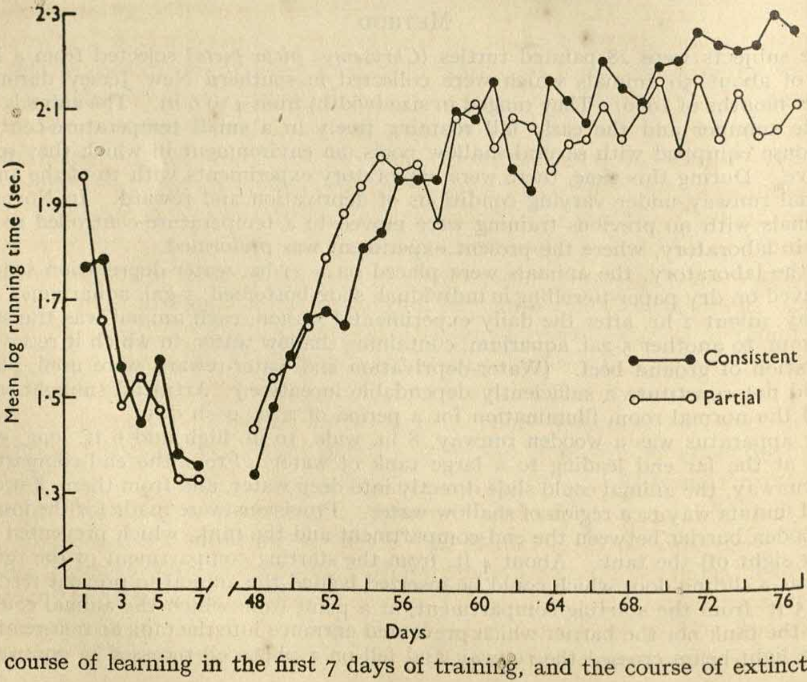
During the first seven days of the experiment, all runs were reinforced. On the basis of their performance in this initial stage of training, some animals were discarded for failure to run, and the remainder were divided into two matched groups, which (after a few losses due to occular infection as the work proceeded) numbered 14 animals each. During the next 40 days, training continued as before for the Consistent Group. For the Partial Group, only 20 of the 40 trials were reinforced, the sequence of reinforced and non-reinforced trials having been constructed according to the rules of Gellermann (1933). If, on any training trial, an animal failed to traverse the alley in 5 min., it was urged forward gently, and a time of 300 sec. was recorded. One Partial animal was treated in this way on Trials 1 and 20, another Partial animal on Trial 20, and a Consistent animal on Trial 14.

Extinction began on Trial 48, the procedure on extinction trials being exactly the same as on non-reinforced training trials with the exception that, during extinction, the animals never were urged to run. Upon entering the end compartment, the animal was detained for a period of 1 min. and then returned to its aquarium. If, on any extinction trial, the animal failed to enter the end-compartment in 5 min., it was removed directly to its aquarium, and a running time of 300 sec. was recorded. Work with any animal was terminated after three successive incomplete trials of this sort. The entire experiment was terminated after 78 days.

RESULTS

In Figure 1, mean log time is plotted for the first seven trials of acquisition and for the 31 trials of extinction. On the first seven trials, which were consistently

FIGURE 1



The course of learning in the first 7 days of training, and the course of extinction.

reinforced for all animals, there was a continuous, negatively accelerated decline in running time. The curves for this period illustrate the degree of control which was achieved over the behaviour of the animals, as well as the equality of the two groups to the point at which partial reinforcement was introduced. During the next 40 days, however, the running times of the Partial Group began to exceed those of the Consistent Group. An analysis of variance based on these times yields a significant Groups effect ($F = 7{\cdot}05$ with 1 and 26 d.f., $p < 0{\cdot}05$), a significant Days effect ($F = 4{\cdot}37$ with 39 and 1,014 d.f., $p < 0{\cdot}05$), and an interaction which falls short of significance ($F = 1{\cdot}09$ with 39 and 1,014 d.f., $p > 0{\cdot}05$).

At the start of extinction, then, there was a difference in the running times of the two groups. That difference, and its rather rapid disappearance, may be seen in the extinction curves of Figure 1. Later in extinction, the two groups began to differ once more, but in the opposite direction; while the Partial Group ran more slowly at the outset, the Consistent Group ran more slowly toward the end. This new divergence of the curves reflects a differential resistance to extinction, greater in the Partial Group than in the Consistent, which began to be manifested only in the last 10 days of extinction. At the termination of the experiment, 10 of the Consistent animals and four of the Partial animals had met the criterion. The probability of obtaining a difference this large by chance is approximately 0·05 by the Median test. The experiment was terminated at this point, because the test could not be affected by the outcome of further trials.

Discussion

These results may be compared with those obtained in analogous experiments with rats (Weinstock, 1954, 1958), with immature turtles of another species (Eskin and Bitterman, 1961), and with African mouthbreeders (Longo and Bitterman, 1960). Each experiment involved a simple instrumental response (running in a runway or striking at a target), a temporal measure of performance, a 24-hr. intertrial interval, and equated trials with 50 per cent. or 100 per cent. reinforcement. The results for rats show greater resistance to extinction after partial than after consistent reinforcement. The results for mouthbreeders show precisely the reverse effect—greater resistance to extinction after consistent than after partial reinforcement. The results for turtles are in a certain sense intermediate between those for the higher and the lower forms. In the experiment with immature sliders, there was no appreciable difference in the resistance to extinction produced by the two percentages of reinforcement. In the present experiment, there was some indication of greater resistance to extinction after partial reinforcement, but it was a small effect, of borderline statistical significance, which appeared only in the later stages of extinction. The effects found in parallel experiments with fish and rat have developed early. In this respect, the present results are reminiscent of those obtained in experiments on the fish under conditions of massed practice (Wodinsky and Bitterman, 1959, 1960); in those experiments, differences between Partial and Consistent Groups, although opposite in direction to the one found here, also appeared only rather late in extinction.

Just how this array of findings should be interpreted, it is difficult to say at present. Resistance to extinction in any given species probably is mediated by a variety of interrelated processes whose relative importance, at the very least, may vary from species to species, if, indeed, the same processes may be assumed to operate in all species. It is conceivable that the factors responsible for the paradoxical effect in the rat tend also to produce such an effect in the turtle, and in the fish, but are overborne by competing factors which are weighted more heavily in the turtle, and still more heavily in the fish. It is conceivable, too, that certain of the factors which make for the paradoxical effect may be unique to higher animals. Much more work may be necessary to give us some fundamental understanding of these problems. A few comparative experiments do, however, seem to have put them in more realistic perspective, dispelling some of the oversimplified notions engendered by rat-centred research.

This study was supported by the Office of Naval Research under terms of Contract Nonr-2829 (01). Mr. Gonzalez participated in the work during his tenure of a Post-doctoral Fellowship awarded by the National Institute of Mental Health, U.S. Public Health Service. Miss Ellen Gower gave valuable assistance in the preliminary stages of the experiment.

REFERENCES

BITTERMAN, M. E. (1960). Toward a comparative psychology of learning. *Amer. Psychologist*, **15**, 704–12.

ESKIN, R. M., and BITTERMAN, M. E. (1961). Partial reinforcement in the turtle. *Quart. J. exp. Psychol.*, **13**, 112–6.

GELLERMANN, L. W. (1933). Chance order of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, **42**, 206–8.

GONZALEZ, R. C., LONGO, N., and BITTERMAN, M. E. (1961). Classical conditioning in the fish: Exploratory studies of partial reinforcement. *J. comp. physiol. Psychol.*, **54**, 452–6.

LONGO, N., and BITTERMAN, M. E. (1960). The effect of partial reinforcement with spaced practice on resistance to extinction in the fish. *J. comp. physiol. Psychol.*, **53**, 169–72.

WEINSTOCK, S. (1954). Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, **47**, 318–22.

WEINSTOCK, S. (1958). Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *J. exp. Psychol.*, **56**, 151–8.

WODINSKY, J., and BITTERMAN, M. E. (1959). Partial reinforcement in the fish. *Amer. J. Psychol.*, **72**, 184–9.

WODINSKY, J., and BITTERMAN, M. E. (1960). Resistance to extinction in the fish after extensive training with partial reinforcement. *Amer. J. Psychol.*, **73**, 429–34.

ENSEMBLE SIZE AND SELECTIVE RESPONSE TIMES WITH A CONSTANT SIGNAL RATE

BY

JOHN BREBNER and IAN GORDON

From the Psychology Laboratory, University of Exeter

Using a selective response task in which subjects were required to respond only to one of a number of alternative digits, although signal rate was held constant, the latency of selective responses was found to lengthen as the number of alternatives increased. This finding does not support the inference drawn from the work of Mowbray (1960) that, with signal rate held constant, there should be no significant lengthening of reaction times as the ensemble of digits increases.

INTRODUCTION

Mowbray (1960), investigating the effects of increasing choice on the speed of the highly practised response of naming Arabic numerals, reports that no increase in vocal reaction times was found when the degree of choice was varied between 2 and 10 signals. These results were obtained in a situation in which each signal had a separate response associated with it. However, in a second situation, which we shall refer to as a "selective response" task, subjects were required to respond only to one of a number of digits and here Mowbray observed that reaction times lengthened as the number of alternatives increased.

In Mowbray's selective response task digits were always presented to the subject at a constant rate of one every 2·5 sec., so that the signal rate, i.e. the frequency of occurrence of digits to be responded to, decreased as the number of alternatives increased. It is this feature of his experiment apparently, which has led Mowbray to claim that, "there is strong, if not conclusive evidence that the variables of inter-stimulus interval and stimulus rate operated to produce latencies that decreased as the interval increased, and increased as the stimulus rate decreased." ("Inter-stimulus interval" refers to the actual intervals between digits to which responses are required.)

Mowbray, in attributing the increase in response latencies to the decrease in signal rate, appears to have dismissed the possibility that the longer reaction times observed are due to the increasing ensemble of digits. Now, in view of his own interesting finding that in choice response tasks the latencies of highly practised responses may be constant over a wide range of signal ensembles, Mowbray's interpretation of his results is an understandable one. Nevertheless, to determine whether the longer response latencies should be ascribed to the increasing number of possible digits or to the decreasing frequency of signals it would be necessary to vary these independently. Accordingly, the present experiment was conducted to test the hypothesis that selective reaction times lengthen as the number of alternative digits increases when signal rate is held constant.

METHOD

Three subjects—the authors and a naïve junior laboratory technician—took part in the experiment. The subjects sat at a table in a darkened room facing, and at a distance of 1 metre from, a one plane digital display unit (Counting Instruments Ltd.) on which illuminated Arabic numerals 1 in. in height appeared at regular intervals and in random order. To avoid the risk of subjects learning the order of digits, a system of two uniselectors was used to provide a pre-selected sequence, taken from a table of random numbers, which repeated in each of the three conditions only after 2,450 digits had been presented. The number of occasions on which particular digits were presented was balanced, as nearly as possible, in the sequence of 2,450 digits. It should be noted, however, that the subjects dealt only with part of the total sequence in any test-session.

In the first condition the digits 1 and 2 were presented to the subject, in the second condition the digits 1, 2, 3 and 4 were used, and in the third and last condition all the digits from 1 to 8 inclusive were shown. In all three conditions the subject responded only to the digit 2. The response required was to depress a morse-type key as quickly as possible with the thumb of the dominant hand. The thumb response was chosen purely because it left the subject's hand in a comfortable, natural position during the experiment. Reaction times were recorded on a chronotron (1,000 millisec. Full Scale Deflection).

Totals of 600 reaction times in each of the three conditions were obtained from the subjects in 18 separate test-sessions in each of which 100 response times were recorded. The order of test-sessions was such that any condition (two, four or eight alternatives) followed all conditions on an equal number of occasions. The most important feature of the experiment is that the rate at which the digits were presented was varied so that the mean interval between signals was held constant. The actual rates employed were 60 digits per minute in the eight-digit condition, 30 per minute in the four-digit condition and 15 per minute in the two-digit condition. Unfortunately, although for the most part each subject had one morning and one afternoon test-session per day, slight irregularities were introduced by unavoidable commitments.

RESULTS

The mean reaction times for the three subjects in the three conditions are shown in Figure 1. The results of analysis of variance carried out on the data are shown in Table I.

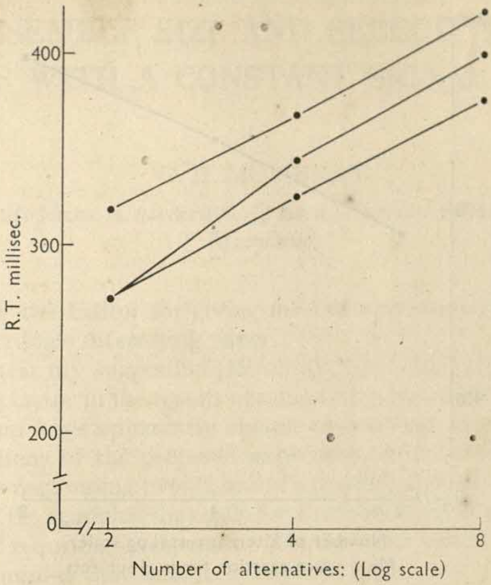
TABLE I

ANALYSIS OF VARIANCE

Source	*d.f.*	*M.S.*	*F. ratio*	*Significance*
Conditions	2	56,538	807·7	$p < 0{\cdot}01$
Subjects	2	10,529	150·4	$p < 0{\cdot}01$
Practice	5	566	8·1	$p < 0{\cdot}01$
Conditions × Subjects ..	4	317	4·5	$p < 0{\cdot}01$
Subjects × Practice	10	167	2·4	$p < 0{\cdot}05$
Conditions × Practice	10	77	1·1	$p > 0{\cdot}05$
Residual	20	70	—	—

While the analysis shows the Practice effect to be significant, it is clear from the Practice × Conditions interaction that the difference between conditions is unaffected by practice, in other words the longer reaction times found with the larger ensembles do not—at least with this amount of practice—become as short as those found with the smaller ensembles. Such a finding confirms Mowbray's own observation regarding selective reaction times. The main finding of the present experiment, however, is that even when signal rate is held constant significant differences are observed in selective reaction times when the number of alternative digits is increased. In view of the extremely large F-ratio found for Conditions the fact that the two interactions Conditions × Subjects and Subjects × Practice are also significant seems relatively unimportant.

FIGURE 1



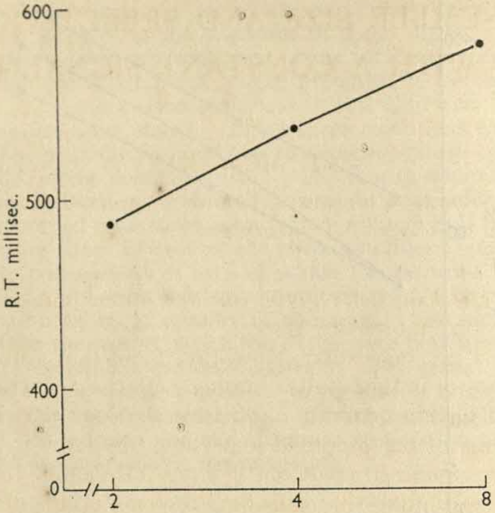
Mean thumb reaction times: 3 subjects

From the finding (Mowbray and Rhoades, 1959) that differences in reaction times, such as are typically observed under varying degrees of choice, can be eliminated with sufficient practice, it could be argued that subjects in the present experiment would, had they been given prolonged practice at the task, have reduced their response times to a constant level. And, although practice appears not to affect the differences between the three conditions in the present experiment, in order to consider the effect of increasing the number of alternatives when the response involved is a highly practised one, it was decided to follow Mowbray's own example in using the highly skilled vocal response of naming digits. In an experiment which paralleled the one described in every respect, except that in this case a voice key was used, and a response was made by calling aloud "Two" whenever the digit 2 appeared, 50 reaction times were obtained from each of four male undergraduates in the three conditions. The only function these results are intended to serve is to indicate whether the relationship between ensemble size and selective response latency, observed by Mowbray for highly practised responses, also holds when signal rate is held constant. The mean reaction times for the three conditions are shown in Figure 2. The difference between mean reaction times proved to be significant ($t = 4\cdot 11$, $p < 0\cdot 05$ between the two and four digits conditions. $t = 3\cdot 77$, $p < 0\cdot 05$ between the four and eight digits conditions).

Discussion

The results presented above support the hypothesis that the effect of increasing the number of alternative digits is to lengthen selective reaction times, and show that the slowing observed by Mowbray cannot be ascribed solely to differences in signal rate. Since, by changing the number of possible digits, we are altering not one, but several variables in the situation, e.g. the proportion of signals and the

FIGURE 2



Number of alternatives: (Log scale)
Mean voice reaction times: 4 subjects

number of categories of digits to which the response is withheld, a more exact specification of what such slowing is to be attributed to must await further research. The relationship between reaction times and the number of alternative digits may, however, be taken tentatively to support the suggestion (Hick, 1952) of a model carrying out a progressive dichotomizing classification resulting in a constant increase in reaction times as the number of alternatives is doubled.

The authors wish to thank Dr. A. C. Staniland for designing and building the apparatus, Mr. D. J. Stone for statistical advice and Mr. W. E. Peckham for acting as a subject in the experiment. They gratefully acknowledge the financial support afforded them by the Nuffield Foundation and D.S.I.R.

REFERENCES

HICK, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

MOWBRAY, G. H., and RHOADES, M. V. (1959). On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

MOWBRAY, G. H. (1960). Choice reaction times for skilled responses. *Quart. J. exp. Psychol.*, **12**, 193–202.

SOME REMARKS CONCERNING BREBNER AND GORDON'S PAPER "ENSEMBLE SIZE AND SELECTIVE RESPONSE TIMES WITH A CONSTANT SIGNAL RATE"*

BY

G. H. MOWBRAY

From The Johns Hopkins University Applied Physics Laboratory, Silver Spring, Maryland

I wish to thank the Editor for giving me the opportunity to see and comment on Brebner and Gordon's interesting paper.

In deciding to test my suggestion (Mowbray, 1960) that stimulus rate may have been an influencing factor in the results obtained with the c-reaction time experiment, Brebner and Gordon have apparently chosen to overlook what I consider to be the important implications of the proposed expectancy hypothesis. In the c-reaction, where the subject is responding to only one of *n* possible stimuli, if *n* is a small number, say two or so, then the stimulus that calls for a response occurs often (usually as often as the stimulus not requiring a response). On the other hand, if *n* is a large number, say 10, then the stimulus that calls for a response occurs rarely relative to those not requiring a response. In the former case, it is my contention that subjects, expecting a response fairly often, maintain an alertness that is lacking in the latter case with a concomitant effect on their reaction times. Furthermore, it is also my contention that within a given choice condition the way in which the stimuli calling for a response are distributed throughout the experimental series will in some measure determine what reaction time will result.

I have collected data, as yet unpublished, that bear out this contention. In a ten-choice c-reaction type of experiment, the key stimuli were distributed in three different ways. In one experimental series the stimuli requiring response were distributed normally, i.e. there were some short intervals and some long ones, but most fell somewhere between; in another, the stimuli were distributed in a rectangular fashion with each interval being represented as often as any other, while in the third condition a "J" distribution was adopted. In this last condition, there were a great many short intervals between responses and progressively fewer long ones. The "J" distribution produced by far the shortest latencies, whereas the rectangular distribution produced the longest ones. This resulted even though all stimuli were represented approximately an equal number of times.

In light of the above considerations, it is not surprising that Brebner and Gordon achieved the results that they did if, as they say, their experimental sequences were randomly constructed. Random selection from among two possibilities leads to distributions of inter-stimulus intervals of the "J" type, whereas selecting from among eight provides distributions that are more nearly rectangular. That is one point. The other is that with only two choices the response/no response ratio is very much higher than with eight choices, assuming equal probability of occurrence of all stimuli. Such is the case no matter what speed of stimulus presentation is used. It is this ratio, I contend, that is important in creating and maintaining expectancy

---

* This paper was prepared under Contract NOrd 7386 between the Bureau of Naval Weapons, Department of the Navy, and The Johns Hopkins University.

on the part of the subject, and it is expectancy that influences reaction times. No more is meant by expectancy than the subjective estimation of the time at which a stimulus calling for a response is likely to occur. If responses occur relatively often with respect to no responses, then expectancies will be high, latencies low, and vice versa.

As further evidence that the ratio of response to no response is important in this type of reaction time, I have another set of unpublished data on a four-choice condition and a 10-choice condition in which the probability of occurrence of the stimulus for response was the same. That is to say, the frequency of occurrence of the response was the same for both four and 10 choices. In this circumstance, the latencies were identical.

In summary, I believe that Brebner and Gordon have, through no fault of their own, misinterpreted my position. My 1960 paper obviously did not make it explicit enough that by stimulus rate I meant the number of stimuli for response relative to those stimuli not requiring a response, and not number of stimuli for response with respect to time.

REFERENCE

MOWBRAY, G. H. (1960). Choice reaction times for skilled responses. *Quart. J. exp. Psychol.*, **12**, 193–202.

APPARATUS

AN APPARATUS FOR RECORDING FIGURAL AFTER-EFFECTS

BY

J. WILSON

From the National Physical Laboratory, Teddington

A simple apparatus is described by which Figural After-effects may be recorded directly from the subject's adjustments of a test figure. Curves which have been recorded confirm the results of other workers obtained by more laborious methods.

Hammer (1949) using the method of adjustment and Ikeda and Obonai (reported by Sagara and Oyama, 1957) using the constant method have measured the time-course of the Figural After-effect, obtaining negatively accelerated curves for the decay of the effect. In this note we describe an apparatus with which such curves can be recorded directly from the subject's adjustments, by linking the adjustable test figure to a pen recorder. Records are easily obtained within a few minutes, providing an effective demonstration and measure of the course of the after-effect.

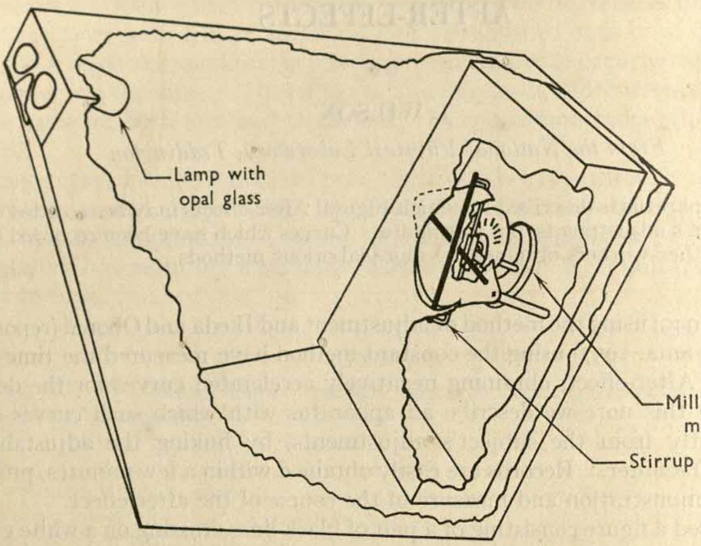
We have used a figure consisting of a pair of black lines crossing on a white ground, the horizontal line fixed and the vertical line movable. The vertical line can be tilted by 15° either way, being fixed by a light stirrup to the movement of a milli-ammeter. Through the stirrup, without touching it, is inserted a card which forms the background of the figure, and on this card is fastened the horizontal line. The lines are made from glass tube, which can easily be drawn out straight and to the required diameter and blackened with drawing ink. Illumination is diffuse and from near the subject's eyes so that no obtrusive shadows are cast on the figure. The meter movement with the stirrup and rod must be carefully balanced and set for zero position at the vertical.

The milli-ammeter which drives the figure is connected in series with the movement of a pen recorder, across a bridge circuit which the subject controls by turning a potentiometer knob, a large knob for sensitivity. The pen thus follows the figure which the subject adjusts by turning the knob. A range switch connects resistors in parallel with the figure, so that full scale of the recorder corresponds to ±16°, ±4°, ±2° of the figure. The time-scale used is 1 in./min.

One of the difficulties encountered was a hysteresis effect in the figuremeter; this was overcome by using a D.C. supply which had not been smoothed. The construction is shown in Figure 1, and the circuit in Figure 2.

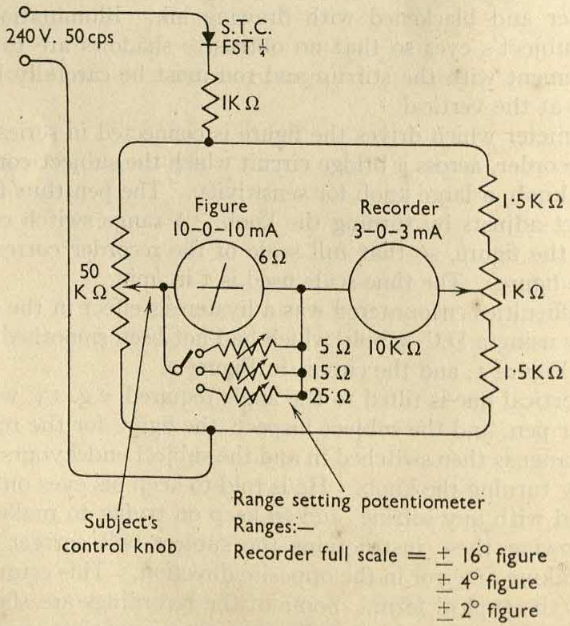
In use, the vertical line is tilted to the angle required, e.g. 13° which can be read from the recorder pen, and the subject inspects the figure for the required time, e.g. 1 min. A finer range is then switched in and the subject endeavours to keep the lines at right angles by turning the knob. He is told to keep his eyes on the intersection, not to be satisfied with any setting, and to keep on trying to make the adjustment perfect. In following these instructions the subject will correct out the Figural After-effect by making an error in the opposite direction. This error will be recorded and is usually of the typical form. Some of the recordings are shown in Figure 3. When shown the curves, the subject will often express surprise, as he was sure that the lines were never far from the right-angle; thus the form of the curve is not consciously produced.

FIGURE 1



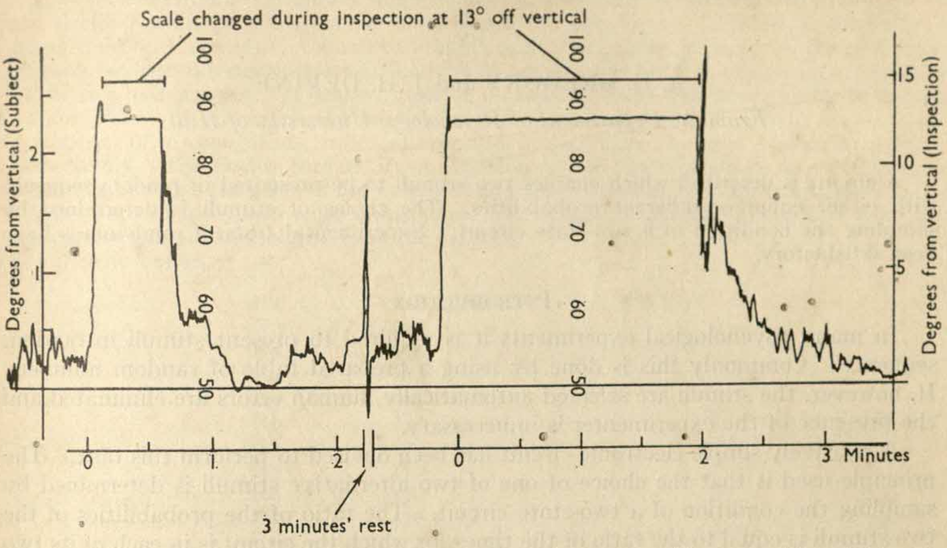
Construction and Mounting of Moving Figure for FAE Recorder.

FIGURE 2



Circuit of FAE Recorder.

FIGURE 3



Recordings of the Figural After-effect. The figure is a cross with an upright which can be tilted. To obtain these recordings the ±16° scale was chosen and the tilt set at 13° for the inspection period. The figure was then switched to the more sensitive scale and the subject required to set and keep the upright at right angles to the cross-bar. The record shows the position during inspection and then the subject's error in setting. Records taken from a naive subject after 5 min. of practice in setting the figure.

The apparatus has been made as part of the research programme of the National Physical Laboratory, and this report is published by permission of the Director of the Laboratory.

REFERENCES

HAMMER, E. R. (1949). Temporal factors in figural after-effects. *Amer. J. Psychol.*, **62**, 337–54.

SAGARA, M., and OYAMA, T. (1957). Experimental studies on figural after-effects in Japan. *Psychol. Bull.*, **54**, 327.

AN ELECTRONIC STIMULUS RANDOMIZER

BY

A. H. GREGORY and J. H. DEVINE

From the Department of Psychology, University of Hull

A circuit is described which enables two stimuli to be presented in random sequence with either equal or different probabilities. The choice of stimuli is determined by sampling the condition of a two-state circuit. Experimental tests of randomness have been satisfactory.

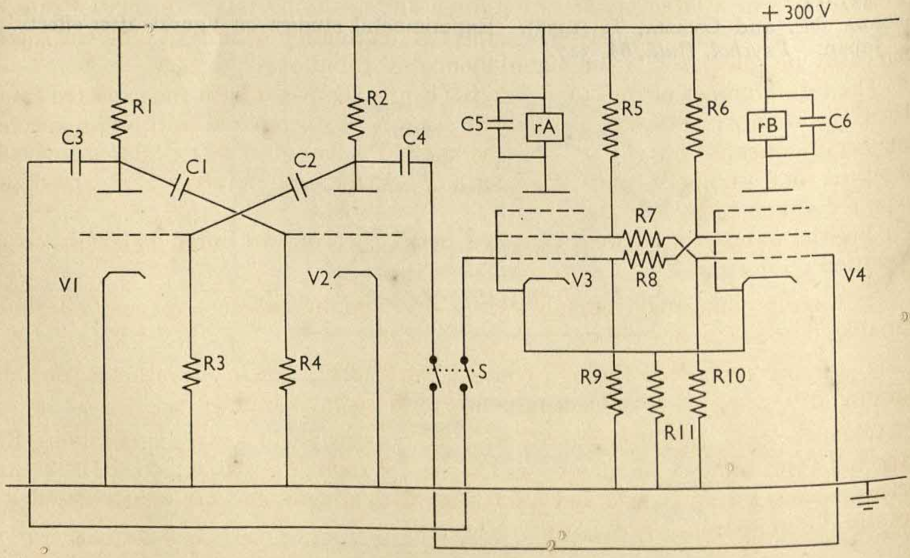
INTRODUCTION

In many psychological experiments it is required to present stimuli in random sequence. Commonly this is done by using a prepared table of random numbers. If, however, the stimuli are selected automatically, human errors are eliminated and the presence of the experimenter is unnecessary.

A relatively simple electronic circuit has been devised to perform this task. The principle used is that the choice of one of two alternative stimuli is determined by sampling the condition of a two-state circuit. The ratio of the probabilities of the two stimuli is equal to the ratio of the times for which the circuit is in each of its two states, i.e. the mark-space ratio.

THE CIRCUIT

The circuit consists of a multivibrator producing continuous square wave oscillations, and a bistable circuit containing two relays which control the stimuli. After the subject has responded to one stimulus, the next one is set by connecting the multivibrator to the bistable circuit for a short interval of time. While connected, the bistable circuit is driven into oscillation by the multivibrator, and on breaking the connection it remains in whichever state it happens to be at that instant. The connection may either be performed by the subject operating a switch between stimuli, or if a response key or switch is used a relay connected to this can be made to close for a short interval. Since the exact interval between successive disconnections depends upon the subject, it will be irregular, and a random sequence of stimuli will result.



The multivibrator is of the symmetrical pattern using a double triode, from which anode pulses are taken at about five cycles per second to drive the control grids of an Eccles-Jordan circuit. If only a single capacitor is used to connect the multivibrator to the Eccles-Jordan circuit, then the multivibrator is unbalanced. Hence pulses must be taken from both anodes. A small padding capacitor in parallel with C1 or C2, or a small variable resistor in series with R3 or R4, may be necessary in order to obtain a final equality of the two half circuits. If desired, resistor R4 may be made variable in order to impose various values of bias upon the relative probabilities of the stimuli. With R4 of value 1 megohm or 1·5 megohms, ratios of stimulus probabilities of about 1 : 2 or 1 : 4 result respectively. The Eccles-Jordan circuit, based upon two power pentodes, is wired in such a manner as to use the grids of the valves for the normal circuit functions, thus leaving the anodes free to supply two 1,000 ohm relays. To avoid any possible distraction from the relays, they should be isolated from the subject.
Component Values.

R1	220 Kilohms	1 Watt
R2	220 ,,	,,
R3	470 ,,	,,
R4	470 ,,	,,
R5	220 ,,	,,
R6	220 ,,	,,
R7	470 ,,	,,
R8	470 ,,	,,
R9	470 ,,	,,
R10	470 ,,	,,
R11	4·7 ,,	3 Watts
C1	0·05 Microfarads	
C2	0·05 ,,	
C3	1·0 ,,	
C4	1·0 ,,	
C5	1·0 ,,	
C6	1·0 ,,	
V1 + V2	ECC 81	
V3	EL 84	
V4	EL 84	

TEST OF RANDOMNESS

A test of the randomness of the stimuli may be made by recording the order of presentations of a large number of stimuli, and analysing the frequency of runs of 1, 2, 3 . . . successive stimuli. The expected distribution of runs of different length can easily be calculated, assuming a binomial distribution.

The total number of runs of all lengths is first compared with the expected total. Then the numbers of runs of different lengths are compared with the expected numbers by means of a χ^2 test. These expected values must be calculated from the observed total number of runs, since for a χ^2 test the total numbers of observed and expected values must be equal.

One thousand stimuli were analysed under each of two different conditions of operation:

(1) Circuit connected through switch operated by subject: on and off times variable, duration about 1 sec.

(2) Circuit connected through contacts of a relay, which closes for a constant duration of 0·7 sec. when subject responds to preceding stimulus.

When the two stimuli, A and B, are equally probable in 1,000 presentations, the expected total number of runs of all lengths for each stimulus is 250. The actual values recorded were (1) 252 and (2) 255 for both stimuli, and are clearly not significantly different from the expected value.

Table I shows the analysis of the frequency of runs of different lengths of the two stimuli under the two conditions.

The observed numbers of runs of length n of the two stimuli, O_A and O_B, are compared with the expected number E, by means of a χ^2 test.

Since the value of χ^2 at the 10 per cent. level is 9·24 for five degrees of freedom, it can be concluded that the stimuli are effectively random.

TABLE I

	n	E	O_A	O_B	$\frac{(O_A-E)^2}{E}$	$\frac{(O_B-E)^2}{E}$
1	1	126·2	125	118	0·011	0·533
	2	63·0	63	70	0·000	0·778
	3	31·4	34	35	0·216	0·413
	4	15·7	17	14	0·108	0·184
	5	7·9	6	11	0·457	1·217
	>5	7·8	7	4	0·082	1·850
		252·0	252	252	$\chi_A^2=0{\cdot}874$	$\chi_B^2=4{\cdot}975$
2	1	127·7	129	120	0·013	0·464
	2	63·7	64	73	0·001	1·358
	3	31·8	29	31	0·246	0·020
	4	15·9	18	21	0·278	1·635
	5	8·0	10	7	0·500	0·125
	>5	7·9	5	3	1·065	3·040
		255·0	255	255	$\chi_A^2=2{\cdot}103$	$\chi_B^2=6{\cdot}642$

BOOK REVIEWS

Brain Mechanisms and Learning. Edited by A. Fessard, R. W. Gerard and J. Konorski. Oxford. Blackwell. 1961. Pp. 702. 70s.

Ours is the age of The Symposium, nourished by generous grants from international foundations, sustained by the willingness of scientists to combine talk and sunshine, and embellished by indiscriminate tape recording. So numerous have been the symposium publications during the past 10 years that one is bound to wonder, upon seeing still another, whether it can really say anything new. In the present instance, the reader will be agreeably surprised. Much of the material will be of fresh interest to psychologist and neurophysiologist alike.

The contributors are as follows: Adey, Anokhin, Asratyan, Buser, Chow, Covian, Doty, Eccles, Eibl-Eibesfeldt, Estable, Fessard, Galambos, García-Austt, Gerard, Grastyán, Hebb, Hernández-Peón, Jouvet, Konorski, Lissák, Luco, Magoun, Morrell, Myers, Náquet, Olds, Palestini, Pinto Hamuy, Rosvold, Segundo and Thorpe.

The contributions can be classified roughly on the basis of their area of major interest: (*a*) psychological, behavioural; (*b*) physiological and anatomical; (*c*) deliberate combinations of (*a*) and (*b*). There is space here only to mention a few studies in each category which struck the reviewer as being of unusual interest.

(*a*) Hebb presents the results of a simple memory experiment which, with a single blow, would seem to dispose of the notion that repetition of digits in a test of memory span involves only "short-term" memory processes, since when a single sequence of digits is repeated from time to time in a larger series of non-repeating sequences, subjects gradually improve in their ability to remember the repeated sequence. "With such results, I can find no way of avoiding the conclusion that a single repetition of a set of digits, with or without reinforcement of being told when an error has been made, produces a structural trace which can be cumulative" (p. 43).

Eibl-Eibesfeldt presents a carefully reasoned argument against various experiments and expositions which have attempted to blur the distinction between innate and acquired patterns of behaviour. In particular she deals with the analysis of the role of experience in nest-building by the rat and in the aggressive behaviour of cats towards rats. Her experiments will be welcomed by many ethologists whose views have been under effective attack by Beach, Lehrman and others.

(*b*) J. and M. E. Olds present a variety of experiments on self-stimulation technique, but one type of experiment is particularly intriguing. By pairing spontaneous electrical discharges in some single units with electrical stimulation to a "reward" point in the medial forebrain bundle, they were able to produce an increase in the rate of firing of the units. Hence, it appears that it is possible to demonstrate operant conditioning of single-unit responses entirely by intraorganic manipulations. Interestingly, neocortical units are not particularly good candidates for reinforcement, whereas paleocortical units often appeared to "learn with ease."

Various authors are concerned with measuring neurophysiological changes which outlast the manipulations producing the changes, such as Eccles and Fessard on post-tetanic potentiation and tetanic potentiation. Morrell deals with an interesting aspect of the spread of an artificially produced epileptogenic focus in one hemisphere to the homologous area in the contralateral hemisphere. By surgically isolating the "mirror" focus after it has become established, he is able to show that its increased irritability is a relatively permanent property and cannot depend on sustained reverberatory activity. "The mirror focus is a region which has not only 'learned' to behave in terms of paroxysmal discharge, but which 'remembers' this behaviour even after months of inactivity" (p. 383). Preliminary observations suggest that changes in RNA distribution are correlated with these alterations.

(*c*) This category is naturally of the greatest interest in discussing a symposium on brain mechanisms *and* learning. There are three ways in which physiological variables are commonly combined with behavioural observations, and all three are dealt with in various papers. These are: (*i*) correlations of electrical recordings and behaviour; (*ii*) the effects on behaviour of stimulation of the nervous system; and (*iii*) the effects of lesions on behaviour.

(*i*) A useful review of his and his collaborators' experiments correlating electrical changes with conditioning is provided by Galambos, who also comments that such correlations are more likely to be concerned with attention and orienting reflexes than with learning as such. Another paper by García-Austt *et al.* applies techniques of recording from the human scalp which, through an averaging procedure, permits the evoked response to be segregated from the background noise level. During the state of attention, it is reported, the evoked potential has a greater amplitude and also is recorded in a much larger area of the brain. These observations clearly relate to earlier animal experiments by Hernández-Peón and others.

(*ii*) Chow reports an experiment in which, in effect, a reversible deficit in visual discrimination learning and retention is produced in monkeys by means of local afterdischarges in the temporal cortex, induced by bipolar square wave stimulation. The results are particularly clear and are almost certainly related (although Chow does not report any control experiments on non-visual discrimination learning) to the modality-specific visual discrimination deficit produced by inferotemporal lesions. It seems evident that the "reversible lesion" technique offers considerable advantages over the irreversible surgical lesion.

(*iii*) Rosvold and Mishkin argue strongly that frontal cortex resections in monkeys produce, in addition to the classical delayed response deficit, an impairment in discrimination tasks carried out under certain conditions. Such an impairment, they claim, is not caused by a change in sensory capacity but by other non-sensory factors, such as a loss of inhibition—although they admit that no single type of disinhibitory effect can adequately account for all the data. This is a paper which, through private circulation, has already had a considerable influence and it is pleasing to see it in print so that one's disagreement or agreement with it can have concrete reference. The lengthy comments of Konorski on this paper will be of particular interest to many readers, since Konorski's laboratory has contributed richly to the analysis of frontal lobe function.

For a final citation which might be of interest to many readers, Hernández-Peón and Brust-Carmona provide a useful review and fresh experimental evidence on the effects of decortication and decerebration on habituation and conditioning, although it is a pity that histological reconstructions are not presented. It would now seem to be beyond dispute that the intactness of the neocortex is not a necessary condition for the occurrence of habituation and conditioning, at least in the cat.

With such a variety of observations, criticism must be directed to individual papers and not to the symposium as a whole, and space obviously does not permit such treatment. One part of the publication which could have been omitted without great loss, with just a few notable exceptions (one of which has already been mentioned) was the *verbatim* account of discussions following each paper. It is clear from this and other symposia that, even from the great, the offhand comment rarely illuminates, and often it is merely puerile.

It is clear that technical skill and methodological sophistication in analysing brain-behaviour relationships are steadily rising. A comparison of this work with research of, say, 25 years ago, seems staggering. And yet the neurological basis of learning remains mercurial. Partly this is due, no doubt, to the complexity of learning processes themselves, so that it is difficult, say, when recording electrical changes during conditioning to know whether they relate to arousal and its habituation or to the establishment of fresh associations; and, of course, to the complexity of the nervous system, vast portions of which are still entirely mysterious. More fundamental, perhaps, is the marked difference in units of analysis appropriate to the two domains—the neurological units being bursts of impulses occurring in a brief period of time, the behavioural units being turns in a maze or a memory of an event that occurred years previously. The operations relating the two are, for the most part, missing, so that research must content itself with either simple correlations—the hippocampus gives 5 to 6 c.p.s. wave-trains as an animal approaches food—or a listing of necessary neurological conditions (rarely has it been possible to state precisely the sufficient conditions) for certain forms of behaviour to occur—the monkey must have two intact frontal lobes to perform delayed response efficiently. Again, it is clear that the major emphasis of much current research is simply misplaced. Whatever else happens in the nervous system when an animal learns, somewhere there must be an enduring structural change which can survive virtually without known alteration in electrical activity as it is conventionally studied.

Despite these difficulties, and possibly because of them, it is clear from a symposium such as this that if there is a genuine frontier in biology, this is one. Frontier life may be rough and insecure, but it is exciting.

L. WEISKRANTZ.

Science Editions. John Wiley & Sons in collaboration with the Basic Book Club, Inc. New York. 1961.

This is a new paperback series concerned with Pure Science and Social Science. It includes the following four titles of interest to readers of this *Journal.*

Organization of Behavior. By D. O. Hebb. Wiley (1949). New York: Science Editions, Inc. 1961. 15s. net. Reviewed in this *Journal,* 1950, **2**, 142.

On Human Communication. By Colin Cherry. Chapman & Hall (1957). New York: Science Editions, Inc. 1961. 15s. net. Reviewed in this *Journal,* 1958, **10**, 111.

Emotions and Memory. By D. Rapaport. Menninger Monograph Series. No. 2 (1942). New York: Science Editions, Inc. 1961. 15s. net.

This monograph provides a comprehensive survey of the clinical and experimental data relating to memory and affective processes. It has not, however, been revised to take account of work reported since original publication.

Comparative Psychology of Mental Development. By Heinz Werner. Harper (1940). New York: Science Editions, Inc. 1961. 19s.

A well-known study of the development of perception and thought from the *Gestalt* standpoint. O. L. ZANGWILL.

Essays presented to Professor Wolfgang Köhler on the occasion of his 75th birthday. Edited by Esther Pabst, S. E. Asch and W. Witte. *Psychologische Beiträge.* 1962. Vol. 6. Parts 3–4. Pp. 707.

This *Festschrift* will be warmly welcomed by experimental psychologists everywhere. Although Gestalt theory is nowadays in the shadows, Professor Köhler is happily very much with us and very much in touch with contemporary ideas. Among the contributors are: G. W. Allport, R. Arnheim, F. C. Bartlett, R. Held, H. Helson, G. Humphrey, I. Kohler, D. Krech, A. R. Luria, A. Michotte, W. D. Neff, H. Wallach, H. Werner and W. Witte. Most papers are in English but with German and French Abstracts. The frontispiece shows a very youthful-looking septuagenarian in relaxed mood. A bibliography of Professor Köhler's writings from 1909 to 1960 is appended.

O. L. ZANGWILL.

The Crisis in Psychiatry and Religion. By O. H. Mowrer. Insight Book No. 1. Princeton and London. D. van Nostrand Co. Inc. 1961. Pp. viii + 264. 15s.

This book brings together 13 lectures and previously published articles which in spite of their varied occasions develop a coherent theme. Any book in this area by an *experimental* psychologist, especially one of Mowrer's background and reputation, is likely to attract attention, and the present volume will not disappoint.

The author points out that some features of psychoanalytic theory which, logically, lead to irresponsibility and libertine behaviour, make it an unsatisfactory guide to the treatment of psychiatric disorders. His alternative, which is to regard mental disorder as an attempt to adapt to a social setting and to treat mental patients as in some measure responsible for their condition, suggests that much could be learnt from the Christian religion. This, after all, with Judaism from the time of the Great Prophets onwards, has been dealing with problems of personality disorder for about as many half centuries as Psychoanalysis has years. Although empirical knowledge built up over a long period may have its shortcomings it is never negligible, and usually proves deserving of considerable respect.

The author recognizes that what he suggests may not apply to all patients and is obviously prepared for the reader to argue with him. His views are, however, undoubtedly challenging and even if they are heretical in terms of current psychological dogma, they have the scientific orthodoxy of staying close to known facts.

The straightforwardness of the writing makes this book suitable for reading by those who have not been trained in psychology or psychiatry. For this reason it should help to end the seduction of clergy by Psychoanalysis which has produced over the past 40 years a series of books whose dubious psychology and immature theology have commended neither discipline to serious students of the other. A. T. WELFORD.

Neuropsychopharmacology. Volume 2. Proceedings of the second meeting of the Collegium Internationale Neuro-psychopharmacologicium, Basle, 1960. Edited by E. Rothlin. Amsterdam. Elsevier. 1961. 521 pp. 120s.

This volume consists mainly of the reports of five symposia held at this meeting. The subjects covered were (i) antagonism to psychotropic drugs; (ii) the effect of psychotropic drugs on conditioned responses; (iii) the influence of specific and non-specific factors of the clinical effects of these drugs; (iv) measurements of their human behavioural effects, and (v) biochemical mechanisms related to their action. Each symposium consists of two main papers followed by a discussion and then a number of shorter individual communications. The quality of these is usually high, and very recent experimental facts and new hypotheses are presented in a manner that will be of the greatest use to anyone seeking to keep abreast of developments in this rapidly developing and many-faceted field. Of particular value are papers on drug antagonism by Gaddum; on the interaction of drugs and behaviour by Cook and Kelleher; on brain monoamines by Carlsson; on biochemical mechanisms and ultrastructure by Richter; a number of papers by Dr. Himwich's group and a sophisticated paper by Shepherd on the complex field of evaluating the clinical effects of these drugs. The countries of origin were also of interest—the U.S.A. having 28 contributors (not counting discussions from the floor); France 14; United Kingdom 13; Italy 7; Iron Curtain countries 5; Canada 3; rest of Europe 13. Thus the U.S.A., while still in the lead, is not altogether dominating the field!

If one then analyses the contribution according to subject—one method of finding out current research trends—it is apparent that interest is still centred very much on the mode of action of reserpine, phenothiazines, imipramine and monoamineoxidase inhibitors, with particular reference to the role of cerebral catechol amines and 5-HT. Although we have by no means reached the stage of any final answers, progress has nevertheless been most encouraging and we now at least know a lot of facts about the action of chlorpromazine (on e.g. phosphorylation and acetylcholine), of reserpine (on brain catechol amines and serotonin in even greater detail), and of 5-HT, 5-HTP, and imipramine and their complex interaction. Less work seems to have been done on the hallucinogenic drugs in comparison, and certainly we seem to know much less about their mode of action. This would suggest that future work should be directed more towards this problem, as the main aim of all this research is still clinical: that is to discover the biochemical and neurophysiological mechanisms in man whose faulty operation leads to the naturally occurring psychoses. These discoveries can come from a study of the mode of action of drugs that induce or aggravate psychotic symptoms as well as of the drug with the opposite effect. Furthermore the former study is more likely to lead to studies that can be carried out clinically. For a positive biochemical effect induced by a psychotomimetic drug could be tested on the metabolic functions of psychotic patients, and it would appear that a systematic and large scale study along such lines would be more worthwhile than the present work based on trying to discover how certain remedies work. For we may find that the latter work mainly by their non-specific action on mechanisms of anxiety, depression and ideation rather than by any direct action on the specific metabolic faults that we suppose to underlie schizophrenia and manic-depressive psychoses. The recent advances in our understanding of the biochemistry of epilepsy, for example, has been gained not by study of the mode of action of anti-convulsants, but of the study of the action of convulsants such as hydrazides, pyridoxine deficiency and aluminia cream on biochemical as well as neurophysiological systems. The time seems ripe then to make a concerted effort to tackle the problem of the basic mechanisms of psychosis. We have at least as many chemical tools as had the workers in epilepsy. The trouble is that this work needs to be carried out in depth as well as on a broad front. For example, it is not sufficient to know that LSD has a particular effect on one or even two particular neurophysiological systems—e.g. the lateral geniculate body, or the afferent inflow to the reticular formation. Its effect on a large number of systems, including cortex, rhinencephalon and diencephalon, using many different methods of analysis, must be known, not only in isolation, but as compared with the effects of many other drugs—both psychotomimetic and non-psychotomimetic; otherwise we cannot disentangle specific, non-specific and irrelevant effects. Similarly we need to know the action of LSD and its relations as well as non-hallucinogenic drugs on a large number of biochemical systems. Work in this field surely now requires a prolonged period of steady data gathering in such a co-ordinated and systematized way rather than basing one's hopes on finding some short-cut to the answers deriving from some fragmentary and most probably irrelevant information.

John Smythies.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIV 1962 Part 3

VISUAL DISCRIMINATION OF OPEN AND CLOSED SHAPES BY RATS. I. TRAINING

BY

N. S. SUTHERLAND, A. E. CARR and J. A. MACKINTOSH

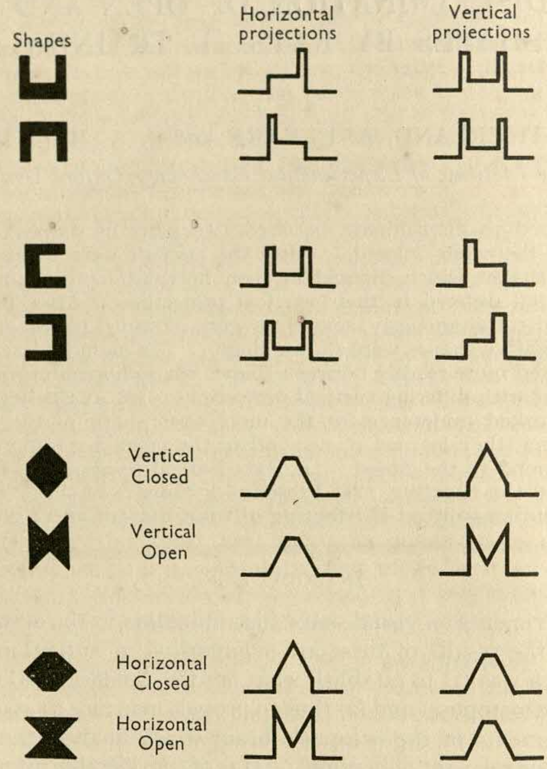
From the Institute of Experimental Psychology, Oxford University

Rats were trained to discriminate between two pairs of shapes: in each pair one shape was "open," the other "closed." Half the animals were trained to discriminate between a pair of shapes which differed in their horizontal projections, the remainder between a pair which differed in their vertical projections. After being trained with both shapes exposed simultaneously, some rats were retrained to respond correctly when the positive and negative shapes were shown singly. The main findings were as follows: (1) Rats discriminated more readily between shapes whose horizontal projections differed than between shapes with differing vertical projections. (2) At the beginning of training rats exhibited a marked preference for the more open shape of the two. (3) During simultaneous training, they learned to respond to the open shape more accurately than they learned to respond to the closed. (4) Rats learned to approach the positive shape rather than to avoid the negative. (5) Presence or absence of shock as punishment for a wrong response and massing or distributing of trials did not affect performance during learning.

INTRODUCTION

A series of experiments on visual shape discrimination in the octopus has recently been undertaken: the results of these are summarized in Sutherland (1960*b*). The purpose of this work was (1) to establish what are the dimensions along which shapes are discriminated by octopuses and (2) thus to provide evidence to assist in unravelling the neural arrangements in the octopus's brain which analyse the input from the shapes in terms of the relevant dimensions. It is of considerable interest to undertake the same kind of work on a second species, both because the question of what type of analysis of shape occurs in other species is important in its own right and because, by having comparative data available, it may ultimately be possible to correlate differences in behaviour with differences in neuroanatomical arrangements: the latter possibility is particularly important since it constitutes an additional method of verifying the function of a given neuroanatomical arrangement. The experiments to be reported in this and the following paper are the first of a series of studies investigating shape discrimination in the rat. In order to explain the design of the present experiment it is necessary first to set out briefly some of the theory which has been evolved to explain results with the octopus. It was suggested (Sutherland, 1957) that one of the shape analysing mechanisms present in *Octopus* operates in the following way: a count of the horizontal extent of the shape is taken at each point on the vertical axis, and a count of the vertical extent at each point on the horizontal axis. The result of counting in the vertical and horizontal directions will be called respectively the vertical projection and the horizontal projection of a shape. It was hypothesized that horizontal extents are more accurately represented in the nervous sytem of *Octopus* than are vertical extents. The reason for suggesting this was that up-down mirror image shapes are more readily discriminable by *Octopus* than are left-right mirror images (Sutherland, 1960*a*), and up-down mirror image shapes differ when counted in a horizontal direction, left-right when counted in a vertical direction. This is illustrated in Figure 1, where the vertical and horizontal projections of two pairs of U shapes (one pair involving up-down reversal, the other right-left reversal) are shown. One prediction which follows from this hypothesis is as follows: if two pairs of shapes are chosen such that the members of one pair differ only in their horizontal projections and the members of the other pair only in their vertical, and the differences in each pair are the same (but appearing on opposite projections), then the pair differing in its horizontal projections should be more readily discriminable than the pair differing in its vertical. The two lower pairs of shapes shown in Figure 1 satisfy these conditions. It was discovered that there was no difference in their discriminability for *Octopus*, and therefore this prediction was falsified (Sutherland, 1960*c*). However, the results of a series of transfer tests suggested that octopuses might be able to perform this discrimination in terms of a second dimension of shape, namely the ratio of contour to square root of area.

FIGURE 1



Horizontal and vertical projections.

The purpose of the present experiments was to investigate the discriminability of these pairs of shapes for rats, and to discover in terms of what dimensions they are discriminated by means of a series of transfer tests. The present paper reports the results of training rats with these shapes, the succeeding paper (Sutherland and Carr, 1962) the results of transfer tests. If the rat analyses shapes in terms of horizontal and vertical extents, and if horizontal extent is more accurately analysed than vertical, then the horizontal pair (shapes 7, 8 in Fig. 1) should be more readily discriminable than the vertical (shapes 5, 6). Other evidence on the rat (Lashley, 1938) suggests that (*a*) rats may be analysing shapes in terms of vertical and horizontal extents since shapes which have the same vertical and horizontal extents (e.g. oblique striations in opposite orientations) are difficult for the rat to discriminate; (*b*) rats have more difficulty discriminating right-left mirror-image shapes than up-down mirror-image shapes, and this might be due to horizontal extents being more accurately represented in the nervous system than vertical extents. The primary purpose of the experiment here reported was to test the prediction that rats would learn to discriminate the horizontal shapes more quickly than the vertical. A secondary purpose was to discover whether performance on the discrimination apparatus used could be improved by spacing trials and introducing shock as a punishment for wrong choices.

Before proceeding to a detailed account of the experiments, it is necessary to say something about the nomenclature which will be used in this and the following paper to describe the shapes. The members of each pair of training shapes besides differing in either their vertical or their horizontal projections differ in other respects. All four shapes have the same area (28 sq. cm.), but they differ in amount of contour (one member of each pair has a perimeter of 32 cm. the other of 20 cm.), and in number of re-entrants, etc. To human beings the first shape in each pair looks more "blobby," more compact than the other: the second shape in each pair looks more "jagged" than the first. In general for shapes of equal area, the greater the length of contour a shape has the more it is likely to contain re-entrants, be judged as "complex," "jagged," and the less likely it is to look compact or "blobby." These characteristics of shapes do not, however, necessarily covary. We shall use the word "open" to describe shapes which have a lot of contour and are jagged, spread-out, etc., and the word "closed" to describe shapes which have little contour, are compact, "blobby," etc. This usage is deliberately vague: if the rat tends to discriminate between open and closed shapes by analysing them in terms of one dimension, we can only give a *precise* definition of "open" and "closed" which will be useful in describing shapes as seen by rats *after* we have been able to specify this dimension by experiment. The four shapes used in training will be named by the names given in Figure 1 (the vertical closed shape, the vertical open, etc.), and these shapes will sometimes be referred to by the letters *VC*, *VO*, *HC*, *HO*.

METHOD

Subjects

The subjects were 32 male hooded rats of a strain originating from the Institute of Dairying Research, Reading. All animals were 4 or 5 months old at the start of the experiment.

Apparatus

The apparatus used has been fully described in a previous paper (Sutherland, 1961). Briefly, it consisted of a start box 8 in. long by 4 in. wide, which was separated from a rectangular choice chamber by a transparent sliding door. The rear wall of the choice chamber was 12 in. from the door leading into it, and was itself 16 in. across. A partition projecting 3·5 in. into the choice chamber divided the rear wall into two sections 8 in. long by 8 in. high. In the geometric centre of each section a drinking nozzle containing milk projected into the choice chamber. The shapes were cut from ¼ in. white Perspex, and had a hole drilled in their centres so that they could be attached to the rear wall, leaving the drinking nozzles projecting through them. The training shapes used measured 8 cm. along their long axis, and 6 cm. along their short. The apparatus was painted matt black except for the wire mesh grid which formed the floor. If an animal approached to within 3·5 in. of the negative shape, an electric circuit was completed which triggered a bell. An A.C. shock of 0·5 milliamp. could be delivered to the animal's mouth, if it attempted to drink from the negative shape. Holding boxes were also present in the experimental room.

Feeding schedule

For 14 days before the start of pretraining, and throughout the duration of the experiment all animals were kept on a 22½-hour hunger schedule. Each group was given *ad libitum* food in the holding boxes in the experimental room at the conclusion of each day's experiment on that group.

Pretraining

On the first day of the experiment animals were placed in groups of four in the experimental apparatus with both milk tubes present, and allowed to drink for an hour. Over the next 7 days, each rat was trained individually to run to its positive shape. The shape was exposed on one side of the apparatus, and a milk tube without a shape was present on the other side. The experimental procedure for all animals was the same as that used for Group I on training trials (*v.* below). Ten trials a day were given, and any animal which failed to run to the positive shape on at least 19 out of the last 20 trials was rejected.

Training

On the next day both shapes were present simultaneously in the apparatus, and discrimination training proper commenced. All animals were run 10 trials a day for 38 days. The side of the positive shape was changed from trial to trial according to Gellerman type sequences: it never remained on the same side for more than three successive trials, and on each day it occupied the left-hand side five times and the right-hand side five times. The procedure on each trial was as follows: the animal was placed in the start box. After 5 sec. the sliding door to the choice chamber was opened and the animal ran to one side or the other. A run to the negative side caused a bell to ring when the animal approached to within 3½ in. of the negative shape. The animal was then lifted out of the apparatus and returned to its holding box. Animals were already used to the bell since this had been in operation during pretraining when animals ran to the opposite side to that on which the single shape was exposed. It was almost invariably possible to pick rats out before they reached the milk protruding through the negative shape, and being picked up did not disturb them. A modified correction procedure was used throughout: i.e. animals which made a mistake were rerun after an interval of about a minute without changing the side of the positive shape. They were rerun in this way until they made a correct choice. In analysing the results only errors made on the first run of each trial were counted. When an animal ran to the correct shape it was allowed 9 sec. drinking the milk before being removed from the apparatus and returned to the holding box.

The training procedure was varied in two ways for Groups I and II. (1) Group I received trials at approximately 2-min. intervals, Group II at approximately 8-min. intervals. Whereas in Group I one subgroup of four animals was run at a time, in Group II all 16 animals were run at once so that Group II animals spent four times as long in the holding box between trials as Group I animals. (2) During training (but not pretraining) Group II received a 0·5 milliamp. shock through the mouth if they attempted to drink from the negative shape. This shock was just sufficient to discourage them from drinking, and after the first 50 trials of training Group II animals almost invariably stopped when the bell sounded and made no attempt to drink from a negative shape.

Retraining on successive discrimination

After the end of training, Group I animals were run on transfer tests for 9 days (*v.* the following paper). On the tenth day, successive discrimination training was commenced. The procedure was the same as during original training except that now only one shape was present on each trial; the milk tube remained in position on the opposite side to that on which the shape was exposed. Animals were rewarded for running to the positive shape, and to the side opposite the negative shape. The bell sounded if they approached the negative shape, or if they ran to the side opposite the positive shape, and animals were then removed from the apparatus. Twelve trials a day were given, on six of which the positive shape was presented and on six the negative shape. Each shape appeared three times on the left and three times on the right on each day.

Experimental design

The shapes on which animals of Groups I and II were trained were exactly the same. Thus the experiment on Group II may be regarded as a replication of the experiment with Group I with the two procedural modifications noted above. There were 16 rats in each of these two groups. The training shapes used were the four lower shapes of Figure 1. From each of the two groups four subgroups of four animals each were formed: an attempt was made to allocate litter mates to different subgroups. The subgroups were as follows: (1) *VC* The vertical closed shape was positive, the vertical open shape negative. (2) *VO* The same two shapes were used, but animals were trained to run to the open shape. (3) *HC* The horizontal closed shape was positive, the horizontal open shape was negative. (4) *HO* The horizontal open shape was positive, the horizontal closed shape negative.

RESULTS

Effect of different training procedures

Table I shows the results of training set out in terms of the percentage of correct runs for different subgroups over the first 200 trials of training (days 1–20) and over trials 201–380 (days 21–38). It is clear that when the experiment was replicated with Group II the results for separate subgroups were almost identical to those obtained from Group I. This means that neither giving a shock on the negative shape nor spacing trials had any effect on learning: it is possible that both variables had an effect, and the two effects cancelled one another out exactly, but this is a rather unlikely supposition. Comparing the four subgroups of Group I with the corresponding subgroups of Group II, the only subgroup which shows any real difference is *HC* (animals trained to select the horizontal closed shape). Over days 21–38 Group I *HC* averaged 91 per cent. correct responses, whereas Group II *HC* averaged 72 per cent. correct. The reason for this is that two animals in the II *HC* group failed to show any evidence of learning. One of these animals lost weight severely during the experiment, and ran much faster than any of the others. Since all other animals had clearly learned by the 380th trial, whereas these two animals were performing at exactly the 50 per cent. level of success it seems safest to exclude them from further analysis. The percentage correct responses for Group II *HC*

TABLE I

RESULTS OF TRAINING

	I				*II*				
Groups	*VC*	*VO*	*HC*	*HO*	*VC*	*VO*	*HC*	(*HC*)	*HO*
Days 1–20	56	66	58	73	56	64	53	(60)	73
Days 21–38	79	81	91	95	85	77	72	(94)	95

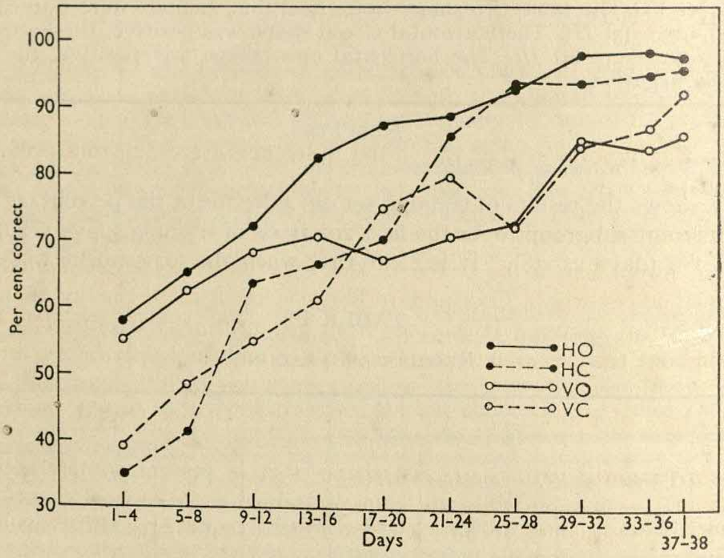
Figures indicate per cent. correct runs.

excluding these two atypical animals is shown in brackets in Table I: it will be seen that the performance of Group II *HC* is now very similar to that of Group I *HC*. The further analysis of training which follows was performed by pooling the results for Group I and II for each subgroup excluding the two animals which failed to learn. None of the conclusions drawn would be affected if these two animals had been included.

Effect of which shape is positive

Figure 2 presents a graph of learning over the 38 days of training for pooled groups excluding the two animals which failed to learn. It is apparent that over the first 40 trials (days 1–4) animals trained to run to a closed shape scored consistently below chance, whereas animals trained to run to an open shape scored consistently above chance. In fact all but one animal of the 16 animals in Groups *VC* and *HC* scored below chance over the first 40 trials: the single animal scoring above chance was one of the two which never learned the discrimination. The probability of 15 out of 16 animals scoring less than 50 per cent. by chance alone is less than 0·001. There can be no doubt that at the start of this experiment animals had a strong pre-existing preference for the open shape: both animals being trained positively and those being trained negatively on the open shape tended to run to it from the very start of training, and this despite the fact that the animals being trained negatively on this shape had already been pretrained to run to the other shape. The large difference between the performance of animals trained positively on open and on closed shapes over the first 200 trials is presumably also due to this preference for the open shape. It will be seen from the *t* tests in Table II that the differences in the performance scores of both Groups *VC* and *VO* and Groups *HC* and *HO* are highly significant over the first 200 trials. Figure 2 and

FIGURE 2



Initial learning.

Table II also reveal that these differences disappear over the last 180 trials, i.e. with sufficient trials the performance of animals trained to run to the closed shapes catches up with that of animals trained to run to the open shapes. Again, this is to be expected on the hypothesis that the initial differences were due to a preference for the open shape.

TABLE II

t TESTS

Groups compared		*Over days*			
		1–20		21–38	
		t	*p*	*t*	*p*
VC	VO	2·8	0·02	0·5	NS
HC	HO	4·9	0·001	1·3	NS
HC	VC	0·2	NS	1·5	NS
HO	VO	3·5	0·01	4·5	0·001

Effect of training on vertical or horizontal shapes

The main hypothesis to be tested in this experiment was that animals trained on horizontal shapes would discriminate between them more readily than animals trained on vertical shapes. It is clear from Figure 2 that Group *HO* performed better throughout the experiment than Group *VO*, and the *t* tests set out in Table II show that the difference between these two groups was highly significant both over the first 20 days of training and over the last 18. To this extent the hypothesis is confirmed.

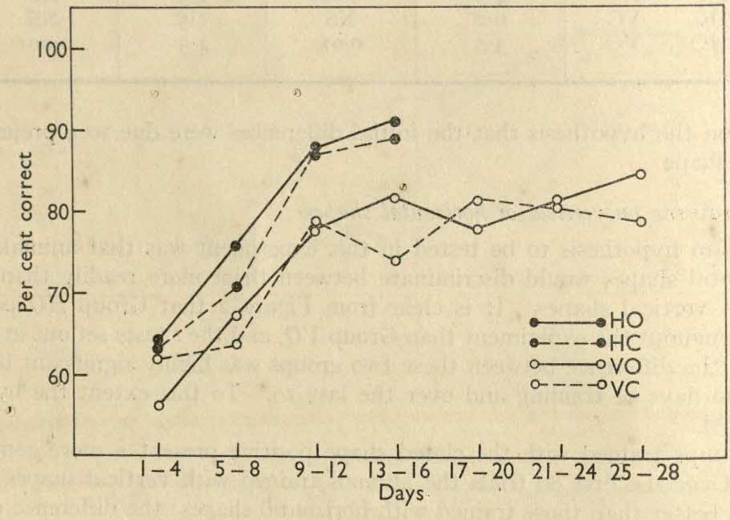
The groups trained with the closed shape positive present a more complicated picture. Over the first 80 trials the animals trained with vertical shapes actually performed better than those trained with horizontal shapes: the difference over this period is significant at better than the 0·05 confidence level ($t = 2{\cdot}4$, d.f. = 12). This is exactly what we would expect if we assume that (1) the open shape is preferred to the closed shape, (2) the horizontal shapes are more readily discriminable than the vertical. For animals being trained with the closed shape positive, the preference for the open shape should show more clearly at first in the animals being trained on the more discriminable pair of shapes: hence Group *HC* made more initial errors than Group *VC*. Thus the worse initial performance of the *HC* group actually confirms the hypothesis that the horizontal shapes are more discriminable than the vertical.

It will be seen from Figure 2, that as learning proceeds Group *HC* comes to perform better than Group *VC* which again we would expect if the horizontal shapes are more readily discriminable than the vertical. However, Table II reveals that the difference over days 21–38, although in the expected direction, is not in fact significant. This failure to find a significant difference between the performance of Groups *HC* and *VC* during the latter stages of training should not lead us to abandon the hypothesis that the horizontal shapes are more readily discriminable than the vertical, since in animals being trained with the closed shape positive there will be a conflict between pre-existing preferences and the effects of training, and this may have increased individual variability. It must be remembered that this hypothesis taken in conjunction with the hypothesis of a pre-existing preference for open shapes explains the following results: (1) The significantly superior performance of *HO* over *VO* throughout the experiment. (2) The significantly inferior performance of *HC* to *VC* at the start of the experiment. (3) The superiority of *HC* to *VC* over the last 18 days of training. (4) The superiority of *HO* to *VO* and of *HC* to *VC* in retraining (*v.* below). It is difficult to see how all these results can be explained except in terms of the two hypotheses outlined above.

Retraining on a successive discrimination

All animals of Group I were retrained to respond correctly when only one shape at a time was shown. Figure 3 gives a graph of performance during retraining: each curve is based on the results of four animals except the *HC* curve which is based on three only since one animal from this group died from a staphylococcic infection in the course of retraining. It will be seen that the animals trained on horizontal shapes received 16 days retraining (192 trials), those trained on vertical shapes 28 days. Group *HO* performed better throughout the first 16 days of training than Group *VO*, and Group *HC* than Group *VC*. Of these differences in performance again only that for the animals trained positively on open shapes is significant. We would, of course, expect these differences to occur, if the horizontal shapes are more readily discriminable than the vertical.

FIGURE 3



Relearning with successive presentation.

Since all groups performed from the beginning of retraining at better than the 50 per cent. level of accuracy, there is clearly considerable transfer from the simultaneous discrimination training they had already received to the same discrimination presented successively. Table III gives a further breakdown of the results for the first 60 and second 60 trials: it is apparent that animals respond much more accurately at first to the positive shape than to the negative shape. This means that at the start of successive training, animals had a very strong tendency to run to the side on which the shape is presented irrespective of which shape it is: thus over the first 5 days of retraining, all groups scored less than 50 per cent. correct on the negative shape, i.e. they ran towards it on more than 50 per cent. of trials. This suggests that animals had learned an approach habit rather than an avoidance habit: the negative shape presumably resembles the positive shape more than does the milk tube exposed on the opposite side of the apparatus to the negative shape, so that when the negative shape was presented alone the animals tended to select it. Subsequent learning is largely a matter of learning not to run to the negative shape, and as animals learn to avoid this the percent runs made to the positive shape actually decreases (compare in Table III the percent correct runs to the positive shape over days 1–5 and days 6–10).

One other interesting feature of the results is that whereas animals trained to run to open shapes perform better when the positive shape is exposed than animals trained to run to closed shapes, they perform worse than animals trained to run to closed shapes when the negative shape is exposed. For example, over the first 5 days Group *HO* score 98 per cent. correct on the positive shape and *HC* 81 per cent.; whereas on the negative shape the *HC* score of 43 per cent. exceeds the *HO* score (32 per cent.). This applies both to animals trained on vertical shapes and to those trained on horizontal, and it holds both for scores over days 1–5 and scores over days 6–10 (*v.* Table III). This means that animals do relatively better with the open shape than with the closed shape, i.e. they do better on the positive shape when it is open than when it is closed, and better on the negative shape when it is open than when it is closed. This suggests that during training the animals had learned more about the open shape than about the closed shape; this tendency is only revealed by comparing scores of animals in different groups, since it is masked in the scores of individual animals by the tendency to score more highly on the positive than on the negative shape.

TABLE III

PER CENT. CORRECT RUNS TO SINGLE SHAPES DURING RETRAINING

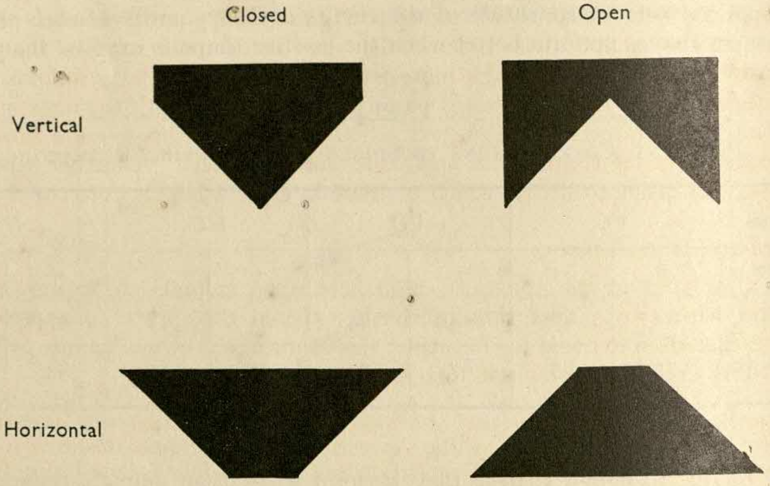
Groups	*VC*		*VO*		*HC*		*HO*	
Shapes	+	−	+	−	+	−	+	−
Days 1–5	82	47	94	25	81	43	98	32
Days 6–10	75	65	87	57	81	81	93	70

DISCUSSION

Dimensions of analysis of shapes

The finding that animals discriminate more readily between shapes whose horizontal projections differ than between shapes whose vertical projections differ is striking. It suggests that as in the octopus so in the rat an analysis of shapes along their x and y co-ordinates may be taking place. We must, however, be careful not to assume at this stage that the same discriminatory mechanisms are at work in the rat as in the octopus: the theory put up to account for shape discrimination in the octopus is based on a considerable quantity of evidence systematically collected, whereas there is little systematic evidence on the way in which rats discriminate between shapes. As mentioned in the introduction two pieces of evidence provided by Lashley (1938) are in line with the theory here being tested. It is worth considering whether the difference in discriminability between the two pairs of shapes here used could be explained in some other way. Rats are known to rely more on the lower halves of shapes than the upper (Munn, 1930; Lashley, 1938), and this has been found to be true to some extent even in the present apparatus (Sutherland, 1961). Figure 4 shows the bottom halves of the two vertical shapes

FIGURE 4



Lower halves of the training shapes.

and of the two horizontal shapes. To human beings there is no question that there is a bigger difference between the lower halves of the vertical shapes than of the horizontal shapes. This does not prove that this is true for rats, but it is suggestive. Thus it seems likely that in terms of their lower halves the vertical shapes should have been more rather than less discriminable than the horizontal. Further evidence will be presented in the following paper that rats were not discriminating between the shapes here used in terms of their lower halves. Thus the evidence set out here supports the hypothesis that one way in which rats analyse shapes is in terms of their vertical and horizontal extents, and makes it worth investigating this hypothesis more fully.

Preferences

The clear-cut pre-existing preference for the open shape is a second very striking finding. This preference may, of course, have resulted from some accidental feature of our subjects' previous experience, but this seems unlikely. Thus the feeding hoppers in animals' home cages were rectangular in shape and resembled the closed shape more than the open. It seems probable that the open shape is somehow more complex to the rat, and therefore more interesting and that the tendency to run to this shape may be a manifestation of some sort of exploratory activity: for another example of rats coming to select the more complex of two shapes, cf. Dember, Earl and Paradise (1957). Preferences for different shapes have been more fully investigated in fishes than in rats, and the usual finding is that fishes prefer more complex, more contoured, more open shapes (cf. Herter, 1929—goldfish; Herter, 1930—perch; Zunini, 1937 and Hager, 1938—minnows): the only exception to this type of finding is Meesters (1940) who found that sticklebacks have a slight preference for less contoured figures. Thus our result with rats is in general agreement with the results on fishes.

During initial training, the rats learned more about the preferred shape (*O*) than about the non-preferred (*C*). This may have been because the preferred shape was the more attention compelling; and, if so, this would represent an interesting and novel finding. However, there is a second possible interpretation of the result which cannot be ruled out. The animals trained with open shape positive may have responded better than those trained with closed shape positive, when presented with the positive shape alone because they had made more runs to the positive shape during training (i.e. had made more correct choices). Similarly, animals trained with *O* negative may have responded better when presented with the negative shape alone than those trained with *C* negative because they had made more runs to the negative shape during training (i.e. had made more mistakes). A further experiment is necessary in order to decide which of these interpretations is correct.

Approach-avoidance learning

It is clear that in the apparatus used here when animals are trained without shock and with two shapes simultaneously exposed they learn to approach the positive rather than to avoid the negative: this confirms a previous finding of Sutherland's (1961) and does not merit further discussion here.

This work was undertaken as part of a project on "Stimulus analysing mechanisms": the project is jointly supported by the American Office of Naval Research (Contract N62558–2453) and by the Nuffield Foundation: we are grateful to these bodies for their financial support.

REFERENCES

DEMBER, W. N., EARL, R. W., and PARADISE, N. (1957). Response by rats to differential stimulus complexity. *J. comp. physiol. Psychol.*, **50**, 514–8.

HAGER, H. J. (1938). Untersuchungen über das optische Differingierungsvermögen der Fische. *Z. vergl. Physiol.*, **26**, 282–302.

HERTER, K. (1929). Dressurversuche an Fischen. *Z. vergl. Physiol.*, **10**, 688–711.

HERTER, K. (1930). Weitere Dressurversuche en Fischen. *Z. vergl. Physiol.*, **11**, 730–48.

LASHLEY, K. S. (1938). The mechanism of vision: XV Preliminary studies of the rat's capacity for detail vision. *J. gen. Psychol.*, **18**, 123–93.

MEESTERS, A. (1940). Uber die Organisation des Gesichsfeldes der Fische. *Z. Tierpsychol.*, **4**, 84–149.

MUNN, N. L. (1930). A note on Lashley's method for studying vision in the rat. *J. genet. Psychol.*, **37**, 528–30.

SUTHERLAND, N. S. (1957). Visual discrimination of orientation and shape by *Octopus*. *Nature*, **179**, 11–3.

SUTHERLAND, N. S. (1960*a*). Visual discrimination of orientation by *Octopus*: Mirror images. *Brit. J. Psychol.*, **51**, 9–18.

SUTHERLAND, N. S. (1960*b*). Theories of shape discrimination in *Octopus*. *Nature*, **186**, 840–4.

SUTHERLAND, N. S. (1960*c*). Visual discrimination of shape by *Octopus*: Open and closed forms. *J. comp. physiol. Psychol.*, **53**, 104–12.

SUTHERLAND, N. S. (1961). Visual discrimination of horizontal and vertical rectangles by rats in a new discrimination training apparatus. *Quart. J. exp. Psychol.* **13**, 117–21.

SUTHERLAND, N. S., and CARR, A. E. (1962). Visual discrimination of open and closed shapes by rats: II. Transfer tests. *Quart. J. exp. Psychol.*, **14**, 140–56.

ZUNINI, G. (1937). Contributo allo studio dell' apprendimento dei pesci, (*Phoxinus laevis* Agas). *Arch. ital. Psicol.*, Fasc. II, E.III, **25**, 3–33.

VISUAL DISCRIMINATION OF OPEN AND CLOSED SHAPES BY RATS. II. TRANSFER TESTS

BY

N. S. SUTHERLAND and A. E. CARR

From the Institute of Experimental Psychology, Oxford University

Rats which had been trained to discriminate between an open and a closed shape were given three series of transfer tests. Two main methods of giving the tests were used: (i) *Successive transfer tests.* Transfer shapes were shown singly, and the animal responded by going towards the shape or to the opposite side of the apparatus to that on which the shape was exposed. In one series of tests using this method, 20 different shapes were shown. (ii) *Simultaneous transfer tests.* Transfer shapes were shown in pairs, and the animal responded by approaching one member of the pair. In one series of tests given by this method, every possible pair which could be drawn from the 20 shapes used in the successive method was presented. On the basis of the tests, it was possible to rank the shapes in order of similarity to the original training shapes for rats. It was found that results obtained by the two methods were consistent: the rank order correlation found between the two rank orders obtained with the different methods was 0·87. The first method was shown to be the more satisfactory since (i) with the simultaneous method position habits tended to mask transfer effects; (ii) fewer trials had to be administered by the successive method to obtain reliable results than by the simultaneous method; (iii) the amount of transfer to each shape with the simultaneous method depends upon which other shapes are included. A high correlation was obtained between the rank order of transfer shapes in terms of their similarity to the original training shapes and their rank order in terms of the ratio $C/\sqrt{A}$ (length of contour divided by the square root of the area). This suggests that rats had learned the original discrimination partly in terms of this ratio: the few results which could not be accounted for in this way could be satisfactorily explained if we assume that the rats were also analysing shapes in terms of their horizontal and vertical projections. It was also shown that rats tend to approach some shapes more than others independently of training. If shapes are ranked according to rats' preferences for them, this ranking also corresponds to that which is obtained by ranking shapes in accordance with the ratio $C/\sqrt{A}$: rats exhibited strong preferences for shapes high on this ratio.

INTRODUCTION

In the preceding paper (Sutherland, Carr and Mackintosh, 1962) evidence is presented which suggests that one way in which rats analyse shapes is in terms of their horizontal and vertical extents. One method of determining the dimensions along which analysis takes place is to present animals trained to discriminate between one pair of shapes with a series of new shapes, and so to discover to what extent each of the new shapes resembles one or other of the originals. The major purpose of the experiments reported in the present paper was to determine the relative similarity for the rat between each of 20 transfer shapes and the training shapes used in the experiment reported in the preceding paper. A secondary purpose was to assess the relative efficacy of two methods of giving transfer tests, one involving the simultaneous presentation of two transfer shapes, and the other the successive presentation of single transfer shapes. If the two original training shapes are discriminated in terms of their position on one dimension we can hypothesize that any other shape will lie somewhere along this dimension. The nearer a given shape lies to one of the original shapes, the greater the extent to which animals will treat it as like that shape. It should be possible to rank shapes in degree of similarity to the originals from the results of transfer tests. In order to do this by the method of simultaneous presentation of transfer shapes it is necessary to present every possible pairing of transfer shapes, otherwise we do not know which of the two shapes presented simultaneously is influencing an animal's choice. This is the first time that this method of paired comparisons has been systematically applied to the problem of shape discrimination, and it is of considerable interest to discover how far the results of this method correlate with the results of presenting transfer shapes one at a time.

METHOD

General procedure

The same animals and apparatus were used as in the preceding experiment on training, and the same notation will be employed here to indicate the different groups and subgroups. To recapitulate: all animals had received simultaneous discrimination training. Group I were trained without shock, Group II with shock. Group I received retraining with the shapes presented successively. Within both Groups I and II there were four subgroups differentiated by whether training was on vertical or horizontal shapes and by whether the closed or open shape was made positive. Thus the subgroups were: (1) *VC* Vertical shapes, closed shape positive. (2) *VO* Vertical shapes, open positive. (3) *HC* Horizontal shapes, closed positive. (4) *HO* Horizontal shapes, open positive. For the four shapes used, *v.* Figure 1 of the preceding paper. Unless otherwise stated below, there were always four animals in each subgroup (I *VC*, I *VO*, I *HC*, etc.).

The procedure for all transfer tests was the same as the procedure used during training, except that the bell and shock were switched off, and animals were always allowed 9 sec. drinking milk whichever side of the apparatus they ran to. Every transfer trial was preceded by a retraining trial in which the procedure was exactly the same as during training. Animals undergoing simultaneous transfer tests were, of course, retrained with the two original training shapes presented simultaneously and animals undergoing successive transfer tests with the original training shapes presented successively. Three separate series of transfer tests were run and are described below: in each series the same transfer shapes were shown to all animals used in that series.

(1) *Simultaneous tests with Group I.* At the end of simultaneous discrimination training, all Group I animals were given nine days of transfer tests with two shapes exposed simultaneously. Ten transfer trials were given on each day, and the pair of transfer shapes used was changed from day to day except that on the first and last day the same pair of transfer shapes was used. Within any one day each member of the pair used appeared five times on the left and five times on the right of the apparatus. In addition, each member appeared five times on the same side as the positive shape on the preceding retraining trial, and five times on the same side as the preceding negative retraining shape.

(2) *Successive tests with Group I.* After retraining on the successive discrimination (*v.* preceding paper), animals of Group I were given 200 transfer test trials with one transfer shape exposed at a time. Twelve trials were given a day, and 20 different transfer shapes were used. Transfer shapes were presented in a random order with the constraint that each shape appeared once in each successive block of 20 trials. Each shape was shown five times on the left and five times on the right of the apparatus. Each shape was presented five times after a preceding retraining trial with a positive shape, and five times after a preceding retraining trial with a negative shape. Each shape appeared on the same side as the preceding positive training shape at least twice, and on the opposite side at least twice; each shape also appeared on the same and the opposite side to the preceding negative training shape at least twice. Group I *VO* now contained only three animals since one had died.

(3) *Simultaneous tests with Group II.* At the conclusion of training Group II animals were given 380 transfer tests with two shapes shown simultaneously. Ten tests were given a day; the same 20 transfer shapes were used as in the successive transfer tests with Group I, and every pair of shapes which can be selected from 20 was presented twice. On the second presentation of each pair of shapes, it was intended to present the members of the pair on opposite sides of the apparatus to the sides they had occupied on the first showing: through a mistake this was not actually achieved for all pairs so that some pairs appeared twice in the same positions. The order in which pairs were shown was random. Group II *HC* contained only two animals.

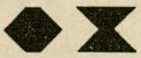
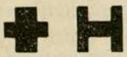
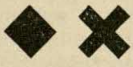
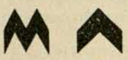
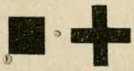
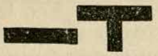
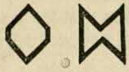
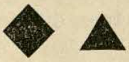
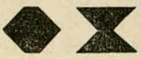
RESULTS

Simultaneous transfer tests: Group I

Table I shows the results of the nine days of simultaneous transfer testing with Group I. For each group the first column of figures indicates the percentage of runs on which the left-hand shape in Table I was treated as equivalent to the original closed shape, i.e. for the two *C* groups the percentage score is based on the number of times the animals ran *to* the left-hand transfer shape, for the two *O* groups it is based on the number of times they ran *away* from it (i.e. to the right-hand shape). A score of 50 per cent. for a given group means that that group performed at chance level during transfer, i.e. they ran an equal number of times to each member of a pair of transfer shapes. Scores of over 50 per cent. indicate that there was a tendency

TABLE I

TRANSFER TESTS: SERIES (1)

	Group	*VC*		*VO*		*HC*		*HO*		
Day	*Transfer Shapes*	*T.T.*	*R.T.*	*T.T.*	*R.T.*	*T.T.*	*R.T.*	*T.T.*	*R.T.*	*Av. T.T.*
1*		70	78	98	90	80	95	90	100	85
2*		68	80	70	80	90	98	83	90	78
3		55	80	83	90	75	88	95	95	77
4		60	85	65	88	50	88	65	98	60
5		48	83	97	93	55	78	88	98	72
6		58	68	80	88	50	73	93	95	70
7*		50	78	38	83	75	80	48	95	53
8		50	70	48	90	60	78	48	90	52
9*		58	85	98	85	50	83	95	88	75

T.T. Per cent. "correct" responses on transfer test trials.
R.T. Per cent. "correct" responses on retraining trials.

All shapes are drawn in the orientation in which they were presented to the *V* Groups: pairs which have a star against them were presented to the *H* Groups at 90° to the orientation shown here.

For fuller explanation, *v.* text.

to treat the left-hand shape in Table I as like the original closed shape (or the right-hand shape as like the original open shape), and scores of less than 50 per cent. indicate that there was a tendency to treat the left-hand shape as like the original open shape In no case does a score of less than 50 per cent. differ significantly from chance expectation. The second column of figures for each group gives the percent correct runs made on each day on the retraining trials with the original shapes.

The same pair of transfer shapes was presented on the first and the last day: it will be seen that the results for Groups *VC*, *VO* and *HO* are very similar on these two days, but there is a big drop in the amount of transfer shown by *HC* from day 1 to day 9. This was largely because the performance of this group was governed by position habits on day 9. The correspondence between the results of the other three groups on these two days suggests that provided position habits do not develop, animals' performance on a given pair of transfer shapes is unaffected by a large number of preceding trials on other transfer shapes. It is also clear that the technique of alternating transfer tests with retraining trials maintains performance on the original discrimination at a fairly constant level.

Examining the detailed results in Table I, it is apparent that wherever a shape not containing re-entrants is paired with a shape containing re-entrants, the former shape is treated as like the closed shape and the latter like the open shape (tests 1, 3, 5, 6, 9) (*v.* preceding paper for discussion of terms "open" and "closed"). Significant transfer occurs in two other cases (tests 2, 4). In (2) the more open shape of the two is treated like the original open shape, and the more closed shape like the original closed shape. The result of transfer test (4) is surprising since here there was a slight tendency to treat the more open shape of the two in the same way as the original closed shape: this will receive further comment below. There is little or no evidence of transfer when two compact, closed shapes are presented (8), nor—more surprisingly—when outline shapes are shown (7).

It is worth noting that with few exceptions the two *O* Groups discriminate between the pairs of transfer shapes better than the corresponding *C* Groups: this is partly to be explained by the superior performance of the *O* Groups on the original discrimination, but that this is probably not the whole explanation will be shown below. The results of the *VO* and *HO* Groups tend to be consistent with one another and also the results of the *VC* and *HC* Groups with the exception of test 7. The explanation of all these findings will be considered in more detail below.

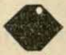
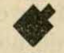
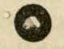
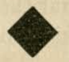
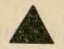
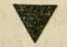
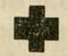
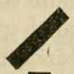
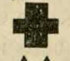
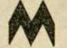
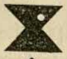
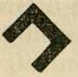
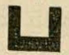
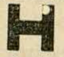
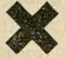
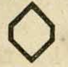
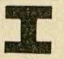
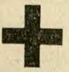
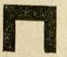
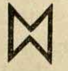
It should be noted that this method of giving transfer tests does not enable us to rank the 16 shapes used in degree of similarity to the original training shapes: nor can we say for any pair of shapes which member of the pair is responsible for producing the transfer. For example, in transfer test 1 we cannot say whether transfer occurs because the rat is treating the rotated closed shape as like the original closed shape, or because it is treating the rotated open shape as like the original open, or for both reasons. The method of giving transfer tests used here is the conventional method, and it was in order to overcome these difficulties in interpretation that the two following methods were developed and applied.

Transfer tests with single shapes: Group I

The results of the series of transfer tests with single shapes are summarised in Table II. This table was arrived at in the following way. If an animal trained to run to a closed shape runs to a transfer shape, it is treating it as equivalent to the original closed shape; if an animal trained to run to the original open shape runs *away* from a transfer shape, it is treating it as equivalent to the original closed shape. We therefore calculated for each of the 20 transfer shapes used the percentage of

TABLE II

TRANSFER TESTS. SERIES (2) AND (3): RESULTS OF SUBGROUPS

	Transfer shapes	*I (Successive)*				*II (Simultaneous)*			
		VC	*VO*	*HC*	*HO*	*VC*	*VO*	*HC*	*HO*
1*		3	2	1·5	4	7	3	3	2
2		1·5	1	6	1	14	4	13	7
3		17·5	3	16·5	2	11	2	11	1
4		7·5	4	14	4	3	1	17	3
5		1·5	5·5	1·5	7	5	10·5	9	4
6		10·5	5·5	14	4	2	6	6·5	13
7		7·5	8·5	9	6	17	8	12	5
8		7·5	10	7·5	8	15	9	6·5	8
9		14	7	14	9	8	13	8	9
10		19	8·5	7·5	10	16	5	19	6
11*		15·5	19·5	4	13	1	16	5	10·5
12		12·5	17	4	18	6	15	10	10·5
13		4·5	12	16·5	18	4	14	16	15
14		15·5	13·5	11·5	15·5	20	7	1	16
15		7·5	17	11·5	13	13	17	4	14
16*		10·5	11	18	13	18	10·5	18	17
17		4·5	15	19	11	9·5	19	20	12
18		12·5	13·5	10	15·5	12	18	2	20
19		17·5	19·5	4	18	9·5	20	14	18
20*		20	17	20	20	19	12	15	19

Figures represent the rank ordering of the shapes: high ranks indicate that transfer shapes were treated like the original closed shape, low ranks that they were treated like the original open.

All shapes are drawn in the orientation in which they were presented to the *V* groups: starred shapes were presented to the *H* groups at 90° to the orientation shown here.

times animals of Groups *VC* and *HC* ran *to* the shape, and the percentage of times animals of Groups *VO* and *HO* ran *away* from the shape. This means that for all groups the higher the percentage score for each transfer shape, the more that shape was treated like the original closed, and the lower the percentage score the more that shape was treated like the original open. Since the percentage scores are difficult to follow in detail, only rank orders are presented in Table II. For each group the shape on which that group scored the highest percentage is ranked 1, the shape on which it scored the next highest percentage 2, etc. This means that where a shape is ranked high for a group, it was being treated as like the closed shape by that group; where it is ranked low, it was being treated as like the open. (The numbers in brackets used in the text refer to the number allotted to each transfer shape in the left-handed column of Tables II and IV.)

It will be remembered that there were four rats in each group except of Group *HC* in which there were only three, since one rat had died in this group during retraining on single shapes. Since there were 10 trials for each rat with each transfer shape, each of the percentage scores was based on a total of 40 trials, except for Group *HC* for which they were based on a total of 30 trials. With this number of trials we cannot expect to obtain a completely reliable ranking of shapes, since even if the number of runs to each shape was ideally spaced out with 40 runs to the first, 38 to the second, 36 to the third, etc., there would only be a difference of two runs for the score on each shape and this would clearly have no significance. Moreover, the whole of the scale is not used: thus Group *VC* ran 36 times to shapes (2) and (5) (the highest ranking shapes for this group), and three times to shape (20) (the lowest ranking shape). Moreover, 16 runs were made to the next lowest ranking shape for that group (shape 10). The ranges are similar for other groups: thus Group *VO* ran away from its highest ranking shape 30 times, and away from its lowest ranking shape no times. One of the reasons why the number of runs given to each shape is not spread out more is that in transfer tests, animals showed a consistent preference for the side on which the shape was exposed. Thus if we summate the number of runs *to* a shape for all shapes for each group, we find that the different groups ran *to* the transfer shape on 70, 76, 73 and 74 per cent. of trials respectively. During the retraining trials the average number of runs to the shape over all groups was 52 per cent. This tendency to run *to* the shape had appeared at the beginning of training on single shapes but had been largely eliminated in the course of training: it reappears, however, during transfer tests. Performance during retraining trials with the original shapes was by no means perfect: the percent correct runs for each group was respectively 83, 77, 79, 86. Since performance on the original training shapes was itself not completely determined by the shapes on which an animal had been trained, we should not expect performance on transfer shapes to depend solely on this, and once again the reliability of the results is correspondingly decreased.

One way of testing the reliability of the results is to compare the ordering of transfer shapes given by animals trained to approach closed shapes with that given by animals trained to approach open shapes. From Table III (row 1), it will be seen that the correlation coefficients thus obtained are small and only one of them is significant. The smallness of these correlations is almost certainly due to the effects of differential preferences for different transfer shapes. The results of original training show that rats have preferences for certain shapes. It is likely that the number of runs to each transfer shape will be determined by the initial attractiveness of that shape for the animal as well as by its similarity to one or other of the shapes used in initial training. Let us suppose (on the basis of the training results) that animals find open shapes more attractive than closed. In the case of animals trained positively on open shapes, preferences will tend to operate in the same direction as the effect of training, i.e. rats will run more often to open shapes both because such shapes are initially more attractive for them and because they have been trained to run to an open shape. In animals trained to run to closed shapes, however, these two factors will work in opposite directions, i.e. their training will tend to make them run away from open shapes, their initial preferences will increase their tendency to run to open shapes relative to closed shapes. The two factors working in opposite directions will result in a reduction in the extent to which different shapes are treated differently and may produce differences in the rank order of the shapes which depend upon initial preferences, not upon training. It is clear from inspection of Table II that this is in fact happening, for example, for both shapes 3 and 4 (circle and diamond) the *VO* group gives a much higher ranking than the *VC*, and the *HO* than the *HC*: this means that all animals have a lower tendency to run towards these shapes than to run towards other shapes—this tendency produces a high ranking for animals trained to approach open shapes, and a low ranking for animals trained to approach closed shapes. Similarly at the other end of the scale we find the opposite tendency at work, thus for shapes (18) and (19) *VO* and *HO* animals yield *lower* rankings than *VC* and *HC* animals: i.e. there was a tendency for all animals to run towards these shapes. For these four shapes the results for Group II are entirely comparable, i.e. shapes (2) and (3) rank higher for *VO* and *HO* animals than for *VC* and *HC*, and shapes (18) and (19) rank lower for *VO* and *HO* rats than for *VC* and *HC*.

TABLE III

RANK ORDER CORRELATIONS (ρ) BETWEEN GROUP SCORES ON TRANSFER TESTS. SERIES (2) AND (3)

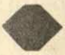
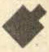
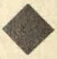
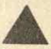
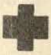
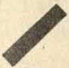
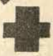
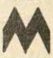
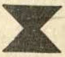
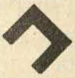
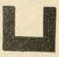
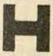
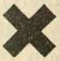
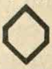
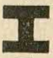
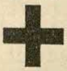
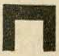
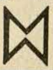
1	I *VC*, I *VO* 0·44*	I *HC*, I *HO* 0·23	II *VC*, II *VO* −0·09	II *HC*, II *HO* −0·05	
2	I *VC*, I *HC* 0·24	I *VO*, I *HO* 0·88***	II *VC*, II *HC* −0·07	II *VO*, II *HO* 0·63**	
3	I *VC*, II *VC* 0·27	I *VO*, II *VO* 0·84***	I *HC*, II *HC* 0·42*	I *HO*, II *HO* 0·79***	
4	I *V*, II′ *V* 0·63**	I *H*, II′ *H* 0·76***	I *V*, II *V* 0·72***	I *H*, II *H* 0·85***	I, II 0·87***
5	I *V*, I *H* 0·74***	II *V*, II *H* 0·80***	*V*, *H* 0·84***		

II′ indicates Group II ranks calculated including trials with side preferences.
* p <0·05.
** p <0·01.
*** p <0·001.

If this interpretation of the results is correct we would expect the correlation between the scores of the *VC* and *HC* groups to be less than the correlation between the *HO* and *VO* groups: for the former groups the two tendencies working in opposition will tend to reduce the spread of percentage scores on different shapes and therefore to reduce reliability. The rank order correlations are shown in row 2 of Table III: though both correlations are positive, the correlation between the scores

TABLE IV

TRANSFER TESTS. SERIES (2) AND (3): COMBINED RESULTS FOR MAJOR GROUPS

1	2	3	4	5	6	7	8	9	10	11
All groups	*Shape*	V_{I}	V_{II}	H_{I}	H_{II}	V	H	$C/\sqrt{A}$	$C/\sqrt{A}$ *ranks*	*Preference ranks*
1*		2	3	2	1	2	1	3·8	2	3
2		3	1	6	3	1	4·5	4·24	4	2
3		1	5	1	8	3	4·5	3·54	1	4
4		5	4	4	2	5	2	4·0	3	1
5		4	7	3	4	6	3	4·56	5·5	7·5
6		6	2	6	12·5	4	9	4·56	5·5	9·5
7		7	8	8	5	7	6·5	4·90	7·5	6
8		9	11	6	7	9	6·5	5·37	9	11
9		8	9·5	10	9	8	10	4·90	7·5	9·5
10		18	6	9	6	11	8	7·6	18	5
11*		17	9·5	11	10	13	11	6·0	13	20
12		15·5	15	12·5	11	16	12	5·69	11	18
13		10	12	17	16	10	16	5·69	11	15
14		15·5	13	16	14	14·5	15	6·53	16·5	13
15		14	18	14·5	12·5	17·5	13	6·04	14·5	15
16*		11·5	14	18	18	12	19	9·20	19	12
17		11·5	17	19	15	14·5	17	6·53	16·5	14
18		13	19	14·5	20	17·5	18	6·04	14·5	17
19		19	20	12·5	17	20	14	5·69	11	19
20*		20	16	20	19	19	20	11·20	20	7·5

Figures represent the rank ordering of the shapes: high ranks indicate that shapes resemble the original closed shape, low that they resemble the original open. Column 9 shows the ratio $C/\sqrt{A}$ for each shape, and column 10 their rank ordering in terms of this ratio. Column 11 gives ranks for shape preferences: high ranks indicate that shapes were approached infrequently, low that they were approached frequently.

All shapes are drawn in the orientation in which they were presented to the *V* groups: starred shapes were presented to the *H* groups at 90° to the orientation shown here.

of animals trained on open shapes is very high and is significant at the 0·001 confidence level, whereas the correlation between the rank scores of animals trained on closed shapes is low and not significant. Exactly the same relationship between the correlation coefficients for the scores of animals in Group II emerges (*v.* right-hand side of row 2 in Table III). Thus the hypothesis outlined above receives very full support.

Since the main object of the transfer tests was to rank shapes in degree of similarity to the two original training shapes, it is necessary to eliminate the effect of preferences from the scores on transfer shapes. The experimental design enables us to do this by combining the scores of animals trained positively and animals trained negatively on the same shape. Thus for Groups *VC* and *VO* the percentage runs to a shape and the percentage runs away from a shape respectively were combined and a new rank order obtained, and the scores of animals of Groups *HC* and *HO* were treated similarly. Since for the *C* groups scores will tend to be depressed for preferred shapes, raised for non-preferred and *vice versa* for the *O* groups, this procedure should tend to eliminate the effect of pre-existing preferences on the transfer test scores. The rank orders obtained after this procedure are shown in Table IV for the group trained on vertical shapes and the group trained on horizontal shapes. There is now a fairly satisfactory correlation between the rankings of transfer shapes for these two groups (I *V* and I *H*, *v.* row 5 of Table III), indicating considerable reliability of the rankings in the two groups.

By considering the percentage scores of the *V* and *H* groups we can determine the point in the series of transfer shapes at which shapes cease being treated as like the original closed and start to be treated as like the original open. Shape (9) was treated as like the original closed shape exactly 50 per cent. of the time by Group I *V*. Since this shape ranked eighth for this group, these animals tended to treat all shapes ranked higher than 8 for them as like the original *closed* shape, and all shapes ranked lower as like the original *open* shape. The crossover point for Group I *H* occurs between the two shapes which rank 9 and 10 for it (these are numbers 10 and 9).

Simultaneous transfer tests: Group II

The results of the transfer tests given to Group II were processed in the same manner as that already described for Group I. It will be remembered that in Group II each of the 20 transfer shapes was twice paired with every other shape. This gives 38 presentations of each shape to each animal. Once again percentage runs to each shape were computed for Groups *VC* and *HC* and percentage runs away from each shape for Groups *VO* and *HO*. There were now four animals in each group, except for Group *HC* which only contained two since the two animals which had failed to learn in this group were eliminated from the experiment. In addition, one animal from Group *VC* had to be withdrawn before it had completed its 380 transfer tests. One difficulty which arose in analysing the data was that it emerged that many animals were developing strong position preferences during transfer tests: as mentioned above, this also occurred in the simultaneous transfer tests with Group I. Since these position preferences are likely to mask transfer effects, scores were calculated in two ways—(i) including all transfer test trials and (ii) excluding transfer test trials on which an animal seemed to be governed by a position habit. It is necessary to adopt an arbitrary criterion of what constitutes a position habit: a position habit was deemed to be influencing the results for a given animal whenever that animal ran to the same side 16 or more times out of 20 successive transfer tests over any 2-day period. Whenever this occurred, the results for that 2-day period were not included under the second method of calculation. Altogether 5,153 transfer test trials were run and in making the calculation excluding runs governed by side preferences 1,657 trials had to be eliminated (33 per cent.). Evidence will be presented below to show that the reliability of the scores obtained excluding runs governed by side preferences was higher than that of the scores calculated on the data including runs influenced by side preferences. Table II presents the rank ordering of the shapes for the different groups based on scores excluding side preferences.

Before going on to discuss these ranks, it is worth giving some idea of how large were the differences in the number of times each shape was selected. Group *VC* selected the shape it ran to most frequently (11) on 69 per cent. of trials with that shape, and the shape it ran to least (14) on 33 per cent. of trials. These figures are based on trials excluding days on which side preferences were shown and since on these trials each shape was shown about 25 times to each rat the total number of trials given was about 100. Percentage scores for the other groups for the shapes most and least approached are about the same *VO* 72 and 32 per cent., *HC* 62 and 35 per cent., *HO* 67 and 30 per cent. It will be noticed that as in Group I the spread of scores is greater for the two *O* groups than for the two *C* groups: this finding has already been explained in terms of the effects of preferences working in the same direction as training for the *O* groups, in the opposite direction for the *C* groups. The *absolute* differences between the number of runs to the most and least selected shape are about the same for the Group II sub-groups as for the Group I, but this is only true because the score for each group on each shape is based on over twice as many trials as for Group I. It should be emphasised again that even with optimal spacing in number of runs to each shape, we would only obtain differences of two or three runs between shapes with neighbouring ranks, and we would thus not expect the ranks to be completely reliable. In terms of *percentage* differences in the number of runs to each shape, differences were smaller for Group II than for Group I: this suggests that for the same number of trials, the method of simultaneous presentation of transfer tests may be less reliable than the method of successive presentation. The smaller percentage spread was not due to Group II animals performing worse during original discrimination training than Group I animals during retraining: indeed, in general they performed rather better; the average number of correct responses made by the different subgroups during retraining was 87, 81, 90, 92 per cent.

It will be shown below that transfer test ranks are about as reliable for Group II as for Group I: however, since nearly twice as many trials were given to Group II as to Group I this suggests that the method of successive presentation yields reliable results with less expenditure of experimenter's time than the method of simultaneous presentation. A further disadvantage of the method of successive presentation is that it is not possible to say where in the list of shapes the crossover point occurs from transfer from the original open shape to transfer from the original closed shape since a score of 50 per cent. for a given shape in simultaneous testing indicates only that if this shape is paired with the other 19 shapes, there is no transfer to it from either of the original shapes. Had the other 19 shapes been different, however, the score on that transfer shape might also have been different.

The first two rows of Table III reveal that when the scores of different subgroups within Group II are correlated, lower rank order correlation coefficients result for each comparison than in the case of Group I. This suggests at first sight that transfer tests with shapes exposed simultaneously are less reliable than tests with shapes exposed singly. However, the insignificant correlations between the rank ordering for animals trained to run to closed shapes and animals trained to run to open shapes revealed in row 1 may merely mean that pre-existing preferences for individual transfer shapes influence performance more when the shapes are exposed simultaneously than when they are exposed successively, and this is supported by further evidence given below. Moreover, when the results are recalculated in such a way as to eliminate the effect of shape preferences (i.e. combining the scores of animals trained on different shapes within the *V* Group and within the *H* Group), rank orders are obtained which have a very high degree of reliability. Thus the correlation between ranks obtained for transfer shapes between Group II animals trained on vertical and on horizontal shapes is extremely high (*v.* Table III, row 5) and this despite the fact that there was zero correlation between the scores of the *VC* and *VO* sub-groups and between the *HC* and *HO* sub-groups. This in itself seems to justify us in our procedure for arriving at transfer test scores purified from the effects of shape preferences, by combining scores for animals trained positively on opposite shapes within the vertical group and within the horizontal group. The reasons for the low correlations between animals trained positively and negatively on the same shapes have already been examined in the last section, and also it was there shown that the higher correlation between *VO* and *HO* than between *VC* and *HC* is fully accounted for by the hypothesis that scores are influenced by two factors—direction of training and pre-existing preferences between shapes. The following considerations lend support to this hypothesis and also explain the correlations shown in row 3 of Table III. Since the scores of animals trained on closed shapes are less reliable than those of animals trained on open shapes because the mutual antagonistic effects of two factors are reducing the spread of scores in the former case, we would expect that correlations between Group I and Group II animals trained to approach closed shapes will be lower than those for Group I and Group II animals trained to approach open shapes. This pattern emerges clearly in row 3 of Table III: there are very high correlations between Group I and Group II animals trained to select open shapes, and small positive correlations between Group I and Group II animals trained to select closed shapes.

A further way of testing the reliability of the two methods of ranking transfer shapes (simultaneous and successive presentation) is to consider the correlations between Group I and Group II animals when the effects of pre-existing preferences have been eliminated by pooling the scores of animals trained positively and negatively on different shapes within the *V* Group and within the *H* Group. These correlations are shown on the right-hand side of row 4 of Table III and are satisfactorily high. The figures on the left-hand side of row 4 give the results of correlating the scores of Group I and Group II animals in this way when the Group II animals' scores are based on all transfer test runs including runs on which there was a side preference. Both for the *V* Groups and for the *H* Groups, the correlations between Group I and Goup II results are lower if Group II scores are calculated including side preference runs than if they are calculated excluding these runs. This proves that the reliability of transfer test scores obtained with simultaneous testing is improved if the results on days on which side preferences were shown are not counted, and it justifies us in presenting in Tables II and IV data on Group II based only on runs not determined by side preferences.

Next it will be seen (Table III, row 5) that the correlation between the results for animals trained on vertical shapes and on horizontal shapes is higher if the ranks yielded by Group I and Group II are combined than if the correlations are calculated separately between the two Group I sub-groups and between the two Group II sub-groups. Ranks were combined by taking the average rank on each shape for I *V* and II *V* and re-ranking to give a combined rank, and similarly for the *H* groups.

The resultant ranks for *V* and *H* animals are shown in Table IV, columns 7 and 8. It is clear that the more groups are pooled, the higher the correlations become, and this indicates that we are obtaining purer measures of transfer effects less contaminated by experimental error.

The last entry in row 4 of Table III indicates the result of combining the ranks given by the *V* and *H* sub-groups for Group I and Group II and measuring the correlation between the Group I ranks thus obtained and the Group II ranks. Once again, ranks were combined by taking the average of the *V* and *H* ranks for each shape, and re-ranking according to the order thus obtained. The resulting correlation of 0·87 is higher than either of the correlations obtained between the I *V* and II *V* Groups and between the corresponding two *H* Groups. The very high correlation produced by combining results in this way indicates that we may place considerable faith in the reliability of pooled results. The order in which shapes are given in both Tables II and IV is the order obtained from combining Group I and II ranks by a repetition of the procedure given above. The interpretation of the final rankings of transfer shapes will be discussed below.

It is worth noting that the final ranking of shapes obtained from the results of the second two series of transfer tests is in complete agreement with the results of the first series of transfer tests. Thus, comparing Table I and Table IV, we would expect good transfer with pairs (1) and (9) of Table I since the rankings of these two shapes (1 and 11 in Table IV) are widely separated. The same applies to pairs (2), (3) and (5) in Table I. For pairs (7) and (8), however, we would expect little transfer because the rankings of the members of these pairs are not widely separated in Table IV. No comparison can be made for pairs (4) and (6) since some of these shapes were not used in the second two series of transfer tests. In every pair where a comparison can be made the left-hand shape in Table I is the one with the higher ranking in Table IV, i.e. the results of the first series of transfer tests and the second two series agree completely about which member of each pair was the more like the original closed shape for rats. The second two series of transfer tests enable us to answer the question which we said could not be answered from the first series. We can now say which shape of each pair is determining transfer: thus in pair (1) of Table I it is the left-hand shape, the rotated closed shape, which produces the transfer since this shape has a very high rank in Table IV indicating that it was treated as very like the original closed shape, whereas the right-hand shape (the rotated open shape) has an intermediate rank in Table IV (11) indicating that there was only a slight rendency to treat it as like the original open.

One further way of analysing the data is of interest. If for all groups we count runs *to* a transfer shape and add up the totals for each shape, this should give us an approximation to the order of initial preferences between shapes. The effects of training are now partialled out since different groups were trained positively on different training shapes: the rankings obtained by this method for Group I and II are not shown separately, but Column 11 in Table IV presents preference ranks obtained by pooling the preference ranks of Group I and Group II animals. Shapes with *high* preference ranks were those run to *least* often, shapes with *low* ranks those run to most often: thus the lower the preference rank for a shape, the more that shape was preferred. Discussion of the detailed preference order will be postponed. It is worth noting here that if we are measuring something different by ranking shapes in order of amount of transfer from the original shapes and in order of preference, we might expect both the correlation between Group I and Group II transfer ranks and that between Group I and Group II preference ranks to be higher than the correlation between Group I transfer ranks and preference ranks and the correlation between Group II transfer and preference ranks. The rank order correlation coefficient between Group I and II training rankings is 0·87, and that between Group I and II preference rankings 0·78, the correlations between Group I preference and training rankings and between Group II preference and training rankings are respectively 0·61 and 0·77, and the correlation between combined training ranks and combined preference ranks is 0·71. Thus it seems likely that we are ordering shapes along different dimensions by the two rankings. There is, of course, still a high correlation between preference and transfer ranks and this is to be expected because there was a clear-cut preference for one of the two original training shapes so that the two dimensions are bound to be correlated.

Discussion

The dimensions along which shapes are discriminated

It was suggested in the preceding paper that rats might be discriminating between the training shapes in terms of their horizontal and vertical projections, and a prediction made on the basis of this hypothesis, namely that animals would discriminate better between the pair of shapes (*H*) differing in their horizontal projections than between the pair of shapes (*V*) differing in their vertical projections, was confirmed. The transfer test results do not support this suggestion. Some of the transfer shapes used were deliberately chosen in order to test for it. The most striking difference in the projections of the two shapes is that the open shape yields two peaks one at either end of the horizontal projection (*H* shapes), or of the vertical projection (*V* shapes), whereas the closed shape yields a single peak in the middle of the projection (cf. Figure 1 in the preceding paper). The two shapes resembling the letter "H" also yield projections having these properties: the upright "H" shape (14) yields two peaks on the vertical projection, the "H" on its side yields two peaks on the horizontal projection. If shapes are discriminated in terms of their projections, then we would expect better transfer from the open shape to the upright "H" for *V* animals than for *H* animals, and better transfer from the open shape to the "H" on its side for *H* animals than for *V* animals. From Table IV (14, 17) it will be seen that *H* animals transferred as well to the upright "H" as *V* animals, and that there was very little difference in the amount of transfer from the open shape to the "H" on its side between the *V* and *H* groups. Similar reasoning applies to two of the "U" shapes (13, 19): both of these shapes on the above hypothesis should resemble the original open shape for *V* animals and the original closed shape for *H* animals. In fact, Group *V* showed less transfer from the open shape to the upright "U" (13) than did Group *H*, and more to the upside down "U" (19). One shape giving two peaks on the opposite projection from the original open shape is, however, not treated as equivalent to the original open shape. This is the 90° rotation of the original open shape (11). This finding—that there is less transfer to the original open shape when rotated than to many other open shapes—is a surprising one, and a tentative explanation will be offered below.

A second possibility which was considered in the previous paper is that animals are discriminating mainly in terms of some feature of the bottom halves of the shapes. If this were correct then *V* animals should tend to treat shapes with a single point at the bottom as like the original closed shape, shapes with a double point as like the original open. There is little evidence that this is happening: thus (Table IV, column 7) the *V* animals treat the cross (18) like an open shape, despite a single narrow bar at the bottom; there is no difference in the treatment of the two "H" shapes (14, 17) although one has a double bar, the other a horizontal line at the bottom; nor is there much difference in the treatment of the two triangles (5, 6) although one has a point at the bottom the other a horizontal line. However, for this group there is a large difference in the treatment of the two "U" shapes—the "U" with the two bars at the bottom (19) resembled the original open shape much more than the "U" with the two bars at the top. The results of the transfer tests also show very clearly that the *H* Group was not discriminating in terms of a long horizontal line at the bottom of the open shape vs. the short horizontal line at the bottom of the closed shape. Thus (Table IV, column 8) shape (5) with a long horizontal base line is treated like the original closed shape, whereas shape (18) with a short base line was treated like the original open.

If animals had been discriminating the original shapes in terms of some feature which would be altered by orientation of the shape, then we would not expect to find a high correlation between the scores of the *V* and *H* Groups since with four exceptions (1, 11, 16, 20) transfer shapes were presented in an orientation which differed by 90° with respect to the orientation of the original training shapes for these two groups. Yet the correlation between the final *V* and *H* rankings was 0·84 and is higher than the correlation between the two subdisivions of the *V* Group (I and II) and the two subdivisions of the *H* Group. Although this high correlation between the *V* and *H* Group ranks has been used above as evidence that transfer test ranks are themselves reliable, it must be remembered that it could not have occurred unless transfer to further shapes was not affected by the orientation of the original shapes to any considerable extent. We conclude, therefore, that the main dimension along which shapes were being analysed was one which is independent of orientation.

In an attempt to pin-point this dimension, we tried out several possible dimensions such as presence or absence of re-entrants. This property obviously gives a poor fit to the result of transfer tests: there is a tendency to treat three shapes containing re-entrants as like the original closed shape (2, 7, 9). Moreover, if shapes were being analysed in terms of the number of re-entrants, it would be impossible to explain why an oblique rectangle should rank consistently lower (i.e. resemble the closed shape less) than a circle (3) or the rotation of the original closed shape (1). One dimension, however, along which the original shapes differ does yield a tolerably good fit to the transfer test results: this dimension is arrived at by taking the ratio of contour to square root of area for each shape ($C/\sqrt{A}$). The ratio is calculated in column 9 of Table IV for each shape, and column 10 shows the result of ordering shapes by allotting them ranks on the basis of this ratio. A correlation coefficient was computed between the transfer ranks of shapes for all groups mentioned (column 1) and the ranks computed from this ratio: its value was found to be 0·87. This high value is all the more remarkable since experimental error must have entered into the final determination of transfer test ranks. Thus the correlation between Group I and Group II rankings is itself only 0·87, and this gives some indication of the extent to which final rankings may have been influenced by experimental error. We thus seem justified in concluding that one of the dimensions in terms of which shapes were being discriminated was that of $C/\sqrt{A}$ or some similar ratio.

One problem in this interpretation is as follows. The original shapes have $C/\sqrt{A}$ ratios of 3·8 and 6·0: some transfer shapes actually have ratios either smaller than that of the original closed shape (3) or considerably larger than that of the original open shape (10, 14, 16, 17, 20), i.e. in terms of this ratio one transfer shape was more "closed" than the original closed shape and five were more "open" than the original open shape. It is not possible to say whether transfer from the open shape should continue to increase as the ratios of transfer shapes increase when these ratios are actually higher than that of the original open shape itself. This would depend upon the extent to which rats transpose along this dimension. Unfortunately, there are not enough shapes of this kind included in the transfer tests to make it possible to decide whether transfer continues to increase as ratios increase beyond that of the transfer shape or whether it then begins to fall off. It is true that the shape with the highest ratio (20) has the lowest rank, but this may partly have been determined by the close resemblance between the vertical and horizontal projections of this shape and those of the original open shape.

If rats were analysing transfer shapes solely in terms of the ratio $C/\sqrt{A}$, certain findings would be left unexplained: these are—(*a*) The difference in discriminability of vertical and horizontal shapes as revealed by the results of training. Such a difference would not be expected if the shapes were being analysed solely in terms of the ratio $C/\sqrt{A}$, since the ratios of the two *V* shapes differ from one another by exactly the same amount as the ratios of the two *H* shapes. (*b*) Transfer to the rotated open shape (11) from the original open shape should be complete: there is no transfer to this shape. (*c*) The "M" shape (10) obtains too high a rank in transfer (i.e. since it has a high ratio, it should be treated as more like the original open shape than it was). (*d*) The upside down "U" shape (19) has too low a rank.

Difficulties (*a*) and (*b*) can be met by supposing that analysis of the shapes was being conducted simultaneously along a second dimension, possibly on the horizontal and vertical projections of the shapes. If the horizontal projections are in some way more accurately analysed than the vertical this would explain (*a*). In analysing the projections, the vertical projection alone will be important for animals trained on *V* shapes because there are no differences in the horizontal projections; similarly, the horizontal projection alone will be of importance for animals trained on *H* shapes. For each group when the open shape is rotated, the pattern produced on the projection relevant to the original discrimination now closely resembles the pattern of the projection of the original *closed* shape. Hence if this method of analysis is being employed together with classification in terms of the ratio $C/\sqrt{A}$, there should be some tendency for the position of the rotated open shape on the transfer rankings to shift nearer to the position of the closed shape than would be predicted from the size of the ratio $C/\sqrt{A}$ above. (*c*) The "M" shape (10) contains finer detail than most of the other shapes: if the rat had poor acuity or if neighbouring contours tend to inhibit one another this would account for its being too near the closed shape on the transfer scale in relation to the size of its $C/\sqrt{A}$ ratio. (*d*) The low position of the upside down "U" shape is more difficult to explain: it should be noted, however, that its position is only too low for the *V* animals, and that for those animals its vertical projection is strikingly like the vertical projection of the original open shape.

These explanations may seem somewhat speculative, but they can be tested: for example, if the high position of shape (11) is to be accounted for in the way described above, we would expected that if animals were trained to discriminate between open and closed shapes such that the open shape gave a very different pattern on the relevant projection when rotated to the pattern yielded by the original closed shape, then the rotation of the open shape should be treated as equivalent to the original open shape, and not as lying midway between the original open and closed shapes. It must be emphasised again that the hypothesis that some such ratio as $C/\sqrt{A}$ is being analysed gives a remarkably good fit to the data, and it is because this fit is so good that the few results which are not explained deserve special attention.

The order of transfer shapes in this experiment is similar to the ordering obtained in a previous experiment with octopuses employing the same pairs of training shapes and some of the transfer shapes used in this experiment (Sutherland, 1960). For example, in the experiment on octopuses, rectangles were treated as less like the closed shape than triangles, the outline closed shape was not treated as equivalent to the original closed shape, etc. The "M" shape was shifted towards the open shape for octopuses as compared to its position in the present results: possibly visual acuity is better in the octopus than in the rat, and therefore the total contour of this shape is taken into account by the octopus's visual analysing system whereas it is not by the rat's.

Shape preferences

It is worth commenting briefly on the rank order of preferences for the different transfer shapes, *v.* column 11 in Table IV: it must be remembered that it is the *low* ranking shapes in this column which are the more strongly preferred shapes. In general, it is the shapes with a high $C/\sqrt{A}$ ratio which are most preferred, but there are exceptions to this. In particular the "M" shape (10) ranks high (i.e. is a non-preferred shape): this supports the hypothesis put up earlier that because of the smallness of the interspaces between its contours this shape may in fact not have all its contours counted in the analysing mechanism in the rat brain. Two other shapes appear in positions which would not be predicted in terms of the ratio $C/\sqrt{A}$, namely the two outline shapes (16, 20). Once again they are too high up the scale. They are obviously very different from the other transfer shapes, and in default of further evidence it is not worth speculating on the explanation of their positions on the preference scale. One very important result is that the rotated open shape is in fact the most preferred shape: it will be remembered that on the transfer scale this shape had too high a rank. Its low rank on the preference scale strengthens the hypothesis that the shapes are being simultaneously analysed along two dimensions: this shape occupies more or less the expected position on the preference scale but it is too high on the transfer scale because of the similarity of its pattern on the relevant projection to the pattern of the original *closed* shape.

Methodology of transfer tests

It has been necessary to comment at length on this topic in presenting the results of transfer tests above. We shall attempt here merely to summarise the principles emerging from the earlier discussion. (1) Since in the rat performance in transfer tests is influenced both by the original training shapes and by pre-existing preferences for shapes, it is essential to train different groups to select opposite members of each pair of training shapes used. (2) If this is done it is possible to obtain highly reliable transfer scores by combining the results of groups trained positively and negatively on the same original shape. (3) The method of giving transfer tests with one shape at a time and the method of giving transfer tests with two shapes simultaneously exposed both produce reliable results, and the results of the two methods are highly consistent with one another. (4) In the situation here used simultaneous transfer testing had the disadvantage that many animals developed side preferences in transfer tests. The reliability of the results is increased if trials on which such preferences were evident are not counted. (5) A further disadvantage of the method of simultaneous presentation is that differences in the performance on different transfer shapes tend to be less than the differences in performance obtained by the method of successive presentation. The reliability of results obtained with simultaneous transfer testing was as high as that of the results of successive as measured by the correlation between the *V* and *H* Groups within each method: since, however, nearly twice as many trials were given in simultaneous tests as in successive, the method of successive tests is more economical in terms of the experimenter's time. (6) The successive method has the further advantage that the amount of transfer to each shape can be measured independently of which other shapes are included in transfer tests: in the simultaneous method, the measure of transfer to each shape will be affected by which other transfer shapes are included.

This work forms part of a project on "Stimulus analysing mechanisms": the project is jointly supported by the American Office of Naval Research (Contract N62558-2453) and by the Nuffield Foundation: we are grateful to these bodies for their financial support.

REFERENCES

SUTHERLAND, N. S. (1960). Visual discrimination of shape by *Octopus*: open and closed forms. *J. comp. physiol. Psychol.*, **53**, 104–12.

SUTHERLAND, N. S., CARR, A. E., and MACKINTOSH, J. A. (1962). Visual discrimination of open and closed shapes by rats. I. Training. *Quart. J. exp. Psychol.* **14**, 129–39.

CHOICE REACTION TIMES AND THE THEORY OF INTERMITTENCY IN HUMAN PERFORMANCE

BY

R. DAVIS

From the Department of Psychology, University of Reading

One visual signal, drawn from two equi-probable alternatives, was followed at a variable interval by a similar signal, also drawn from two equi-probable alternatives. The relationship between reaction time to the second signal and the interval between signals was studied under the following conditions: (1) Subjects were trained from the beginning to regard the first signal as irrelevant. (2) Subjects were asked to report on the nature of the first signal after they had responded to the second. The interval between signals was one of the five values, 50, 100, 150, 200, 250 millisec. In the *regular interval* situation the same interval was used over a block of 20 trials. In the *random interval* situation the values of the interval were randomly arranged, with equal frequency within each block of 20 trials.

The results showed: (1) In the random interval situation there was a definite disturbing effect of the first signal on the reaction time to the second, despite the fact that subjects were trained to disregard the first signal. However, the delays in reaction time were significant only at the shortest interval, 50 millisec. (2) In the regular interval situation this disturbing effect was not evident. (3) The additional requirement of having to report on the nature of the first signal produced no consistent increase in reaction time to the second signal. Some implications of these results for the conception of the human operator as a channel of limited capacity are discussed.

INTRODUCTION

In a series of papers (Davis, 1956, 1957, 1959) the author has examined the conception of the human operator as a single channel information system, after the manner suggested by Craik (1947, 1948), Hick (1948) and Welford (1952), and extended the discussion to the way in which information arriving via different sense modalities is treated. The general design of all these experiments was similar, and consisted of presenting subjects with a pair of signals, separated by a randomly varied interval, and measuring the way in which the reaction time (RT) to the second signal varied as a function of this interval. Using as evidence the relative delays observed in these RTs, it was concluded that subjects did behave in an intermittent manner, and that the period of intermittency was of the order of 200 millisec. In the most recent paper (Davis, 1959) it was argued that paying attention to a signal, rather than performing any overt response to it, was the crucial factor which gave rise to delays in responses to subsequent signals. Fraisse (1957) has produced similar evidence.

These experiments were all performed using "simple reaction time." There was no uncertainty about which signal would arrive, or what the response to it would be. The subject was always informed of the order in which signals would occur; the only uncertainty lay in the exact time at which they happened.

Experimenters using choice reaction techniques, in which subjects did not know in advance which of two signals would occur, have produced contradictory results. Thus, Elithorn and Lawrence (1955), and more recently Halliday, Kerr and Elithorn (1960), have argued on the basis of their results that there is no period of intermittency. Marrill (1957) found that RTs to the second of two signals became longer as the interval between them decreased, but that the RT to the second signal did not usually become longer than the RT to the first signal, even at the shortest intervals. On the other hand, Welford (1959) found data in a two-choice situation consistent with an intermittency theory.

A common feature in the design of all these choice reaction experiments, with the exception of Welford's (1959), is that, although subjects did not know which of two alternative signals would occur first, the second signal was always the opposite alternative. Hence, as soon as the first signal arrived, the second signal was also determined, and consequently there was no longer a genuine two-choice situation for the second signal. It is therefore difficult to know what baseline should be used to assess whether the second RT is delayed. If it is compared with the RT to the first signal this may involve an illegitimate comparison of a simple RT with a two-choice RT.

A number of other issues also need resolving.

First, the possibility existed in one of the author's previous experiments (Davis, 1959), that subjects were having to inhibit a previously learned response to the first of two signals. This could have caused the observed delays. In the present experiment subjects were trained from the beginning to regard the first signal as irrelevant.

The second issue is whether the delay in RT represents a true limit on a subject's capacity. Is his attention wholly devoted to dealing with whatever information the first signal conveys, or could he, at the same time, be coping with other information requiring some form of attention? Of course, the human organism is continuously dealing, more or less automatically, with information, such as the proprioceptive information which regulates posture, and upon which accurate movement depends. The single channel conception is only feasible when applied to man's ability to deal with information which does require some form of attention.

The present experiment attempts to examine this issue by training subjects to regard the first of two signals as irrelevant, and measuring their RTs to a second signal given at different intervals afterwards. These results then serve as a baseline for a comparison in which the same subjects are asked to report on the nature of the first signal, i.e. which out of a number of alternatives it was, after they have responded to the second signal. The purpose is to see whether this aspect of the first signal can be noticed, and subsequently reported, without further delaying the response to the second signal.

Thirdly, by using a multichoice situation for the second signal, as well as for the first, a further possibility can be investigated. One can compare the effect of a regular interval between the two signals with the effect of an interval varied at random, without risking anticipation of the second signal, which tends to occur with regular intervals, when simple RTs are used.

METHOD

An experiment to resolve these issues was designed as follows. It was decided to use the simplest level of multichoice situation to begin with; two-choice, both for interpolated and for response signal.

Apparatus

Four small neon tubes (Ne. 2) were mounted in a square of side 3 in. and were placed orthogonally to the subject's line of sight about 30 in. from his eyes. The top pair of tubes constituted Display 1, the bottom pair Display 2. Each display provided one of two equi-probable alternative signals (left or right) and the time interval between a signal on Display 1 and a signal on Display 2 could be varied. The response panel consisted of two sensitive buttons 2 in. apart. The left button was operated by the first finger of the subject's left hand, the right button by the first finger of his right hand. The response was always to press the left button if the left light in Display 2 came on, and to press the right button if the right light in Display 2 came on. RTs to the nearest millisec. were recorded using a Venner millisec. clock.

Procedure

Each subject went through the following experimental procedure. Normally sessions were held on consecutive days. Each session involved 100 responses.

Session 1. Two-choice reactions were made to the second display only. After a "Ready" warning, a switch was thrown with an audible click, as a further warning signal, and 1·4 sec. after this click, one of the lights in Display 2 came on, and stayed on, until the subject pressed the appropriate button. The left-right occurrence of the lights was randomised with equal frequencies over blocks of 20 trials. The subject was merely told that the left-right sequence was random, and that there was no point in trying to guess which light would come on next. Display 1 was not used at all.

Sessions 2 *and* 3. The subject was told that he would receive a signal on Display 1, shortly before the signal he was to react to on Display 2, but that this first signal was to be disregarded completely, and he had, as before, to respond, as rapidly as possible, only to Display 2. The left-right arrangements of signals on Displays 1 and 2 were randomized independently, with equal frequencies over blocks of 20. The signal on Display 1 occurred 1·4 sec. after the warning click, and lasted for approximately 35 millisec. The signal on Display 2 occurred at a variable interval between 50 and 250 millisec. after the onset of the signal on Display 1, according to the scheme given below. Session 3 was the same as Session 2, but with a different random arrangement of signals.

Sessions 4 *and* 5. Signals were presented in the same way as in Sessions 2 and 3, and the subject was told to respond as rapidly as possible to the second signal. It was emphasized that this was his main task. In addition, he was asked to report, verbally, after he had made the appropriate response to Display 2, whether the signal on Display 1 had occurred on the left or right. There was no pressure on the subject to produce this report as rapidly as possible, although it was normally given without hesitation after the button response. Session 5 was the same as Session 4, but with a different random arrangement of signals.

Three subjects performed under a *regular interval* situation throughout, and three subjects performed under a *random interval* situation. In the *regular interval* situation each of Sessions 2 to 5 was split up into blocks of 20 presentations. Within each of these blocks the interval between first and second signal was constant. Sessions 2 and 4 consisted of five blocks of 20 presentations, with intervals of 250, 200, 150, 100 and 50 millisec., in successive blocks, whereas Sessions 3 and 5 consisted of successive blocks of 20 with intervals of 50, 100, 150, 200, 250 millisec. About 15 sec. elapsed between each presentation, and there was a rest pause of about 1 min. at the end of each block of 20. Thus, although the subject was not told that the intervals were regular, or that they were changed between blocks, each change was marked by a clear break.

In the *random interval* situation, the five values of intervals were randomized so as to occur with equal frequency within each block of 20. Otherwise, conditions were the same.

Subjects. Five of the subjects were undergraduate men at The University of Michigan, between the ages of 17 and 21. The sixth was a woman secretary of 33. All subjects were paid for their services.

RESULTS AND DISCUSSION

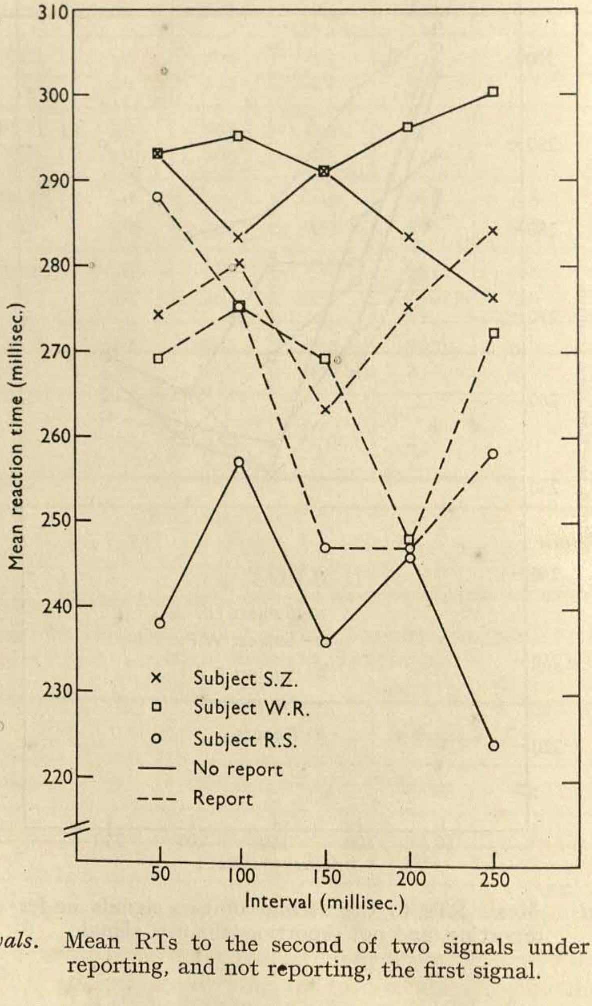
The results for each subject who performed under the random interval situation are shown in Table I, and for those who performed under the regular interval situation in Table II. The "no report" condition consists of the results from Sessions 2 and 3, whereas the "report" condition consists of the results from Sessions 4 and 5. The results for the control two-choice situation (Session 1) are also shown.

From these tables the mean RTs at each interval for each subject under the random interval condition are plotted in Figure 1, and the mean RTs at each interval for each subject under the regular interval condition are plotted in Figure 2. One notices immediately the difference between the systematic variation of RT with interval in Figure 1 as compared with the apparently haphazard variation in Figure 2.

have almost the same mean control two-choice RT (254 millisec. and 253 millisec. respectively).

It is thus evident that subject R.B. contributes most of the interaction (Subjects × Conditions) In Table III. In some respects this was an atypical subject, first, being a woman, secondly, being about 15 years older than the average for the other

FIGURE 2



Regular intervals. Mean RTs to the second of two signals under conditions of reporting, and not reporting, the first signal.

subjects. It is possible that age is a relevant variable in this type of experiment. At each interval, the mean RT for this subject was consistently slightly longer in the "report" condition than in the "no-report" condition. However, a *t* test showed that none of these differences reach significance at $p = 0{\cdot}05$.

Thus, with regard to the effect of conditions, we may conclude that introducing the requirement of reporting the first signal had no significant influence on the RT to the second signal, or at least, it is perfectly possible for a subject to respond just as rapidly in the "report" condition, as in the "no-report" condition.

The largest effect is due to the intervals between signals. This effect may be examined further by considering the magnitude of delays in RT to the second signal, where "delay" is the difference between the mean RT at a particular interval and the control two-choice RT.

TABLE III

ANALYSIS OF VARIANCE FOR THE RANDOM INTERVAL SITUATION

Source	*d.f.*	*MS.*	*F.*	
Subjects	2	18,341	8·51	$p < 0{\cdot}001$
Intervals	4	51,183	23·74	$p < 0{\cdot}001$
Conditions	1	6,252	2·90	—
Subjects × Intervals	8	2,706	1·26	—
Subjects × Conditions	2	23,804	11·04	$p < 0{\cdot}001$
Intervals × Conditions	4	329	<1	—
Subjects × Intervals × Conditions ..	8	1,059	<1	—
Error	1,170	2,156	—	—
Total	1,199			

TABLE IV

ANALYSIS OF VARIANCE FOR THE RANDOM INTERVAL SITUATION, SUBJECT R.B. EXCLUDED

Source	*d.f.*	*MS.*	*F.*	
Subjects	1	12,442	6·44	$p < 0{\cdot}025$
Intervals	4	40,126	20·78	$p < 0{\cdot}001$
Conditions	1	3,771	1·95	—
Subjects × Intervals	4	2,247	1·16	—
Subjects × Conditions	1	3	<1	—
Intervals × Conditions	4	258	<1	—
Subjects × Intervals × Conditions ..	4	1,286	<1	—
Error	780	1,931	—	—
Total	799			

From the "no-report" condition in Table I one may observe:—

(1) For J.F. the delay falls from a maximum of 44 millisec. at interval 50 millisec. to a minimum of 1 millisec. at an interval 150 millisec. A *t* test for this difference gives $p < 0{\cdot}025$.

(2) For W.P. the delay falls from a maximum of 48 millisec. at interval 50 millisec. to a minimum of 8 millisec. at interval 200 millisec. ($p < 0{\cdot}025$).

(3) For R.B. the delay falls from a maximum of 25 millisec. at interval 50 millisec. to a minimum of −12 millisec. at interval 150 millisec. ($p < 0{\cdot}025$).

Thus, as the interval between signals becomes longer, although the decrease in delay is not regular, and although individual differences exist between subjects, there is an overall tendency for a decrease from a maximum delay at the shortest interval of 50 millisec. to a minimum, which does not differ significantly from the control two-choice RT. Although this minimum occurs at 150 millisec. for two subjects and at 200 millisec. for the third, it should be pointed out that the delays are not large enough to reach significance for intervals greater than 50 millisec. In comparing these results with previous work by the author, one might suggest that, although paying attention to a signal, for whatever reason, causes some sort of central delay in dealing with subsequent signals, this delay is not necessarily as long as if an overt response had to be organized to the first signal. One should remark that the delays in the present experiment are measured using a two-choice situation, whereas in the author's previous experiments simple RT was used. However, Welford (1959), using a situation in which an immediate response was required to each of two consecutive signals, each drawn from two alternatives, found delays comparable with those in a simple RT situation.

Regular interval situation

Table V is a summary of an analysis of variance for the regular interval situation. The triple interaction (Subjects × Intervals × Conditions) is significant (F = 2·28, d.f. = 8/1170, $p < 0{\cdot}025$), so that it makes little sense to discuss the main effects in isolation. However, it is clear that the effect due to intervals, is very much less than that in the random interval situation. The largest variance is between subjects, although it should be pointed out that this can be reduced by adding a constant to each RT for subject R.S., to compensate for the fact that the control two-choice RT for this subject is considerably shorter than the control values for the other subject. Thus, if one adds 42 millisec. to each RT for this subject the MS value for subjects in Table V reduces to 23489 (F = 13·17, d.f. = 2/1170, $p < 0{\cdot}001$), which, however, is still the largest main effect.

TABLE V

ANALYSIS OF VARIANCE FOR THE REGULAR INTERVAL SITUATION

Source	*d.f.*	*MS.*	*F.*	
Subjects	2	110,093	61·71	$p < 0{\cdot}001$
Intervals	4	7,426	4·06	$p < 0{\cdot}005$
Conditions	1	7,833	4·39	$p < 0{\cdot}025$
Subjects × Intervals	8	3,793	2·13	$p < 0{\cdot}05$
Subjects × Conditions	2	67,068	37·59	$p < 0{\cdot}001$
Intervals × Conditions	4	5,766	3·23	$p < 0{\cdot}025$
Subjects × Intervals × Conditions ..	8	4,060	2·28	$p < 0{\cdot}025$
Error	1,170	1,784	—	—
Total	1,199			

If the mean RTs at each interval in Table II are compared with the control two-choice RTs, subjects S.Z. and W.R. show no significant variation over the whole range of intervals and conditions. Subject R.S. shows no significant variation over intervals in the "no-report" condition, but in the "report" condition a *t* test shows significant delays for this subject, at the two shortest intervals (55 millisec. delay at interval 50 millisec., $p < 0{\cdot}01$, and 42 millisec. delay at interval 100 millisec., $p < 0{\cdot}025$). However, an exception of this kind does not invalidate the general argument, that it is *possible* to report on the first signal, without delaying the response to the second. The fact that it is possible is clearly demonstrated by the other subjects, in both the regular and random interval situations.

It might be argued that, since subjects always did the "report" condition after the "no-report" condition, if there were an improvement in successive sessions due to practice, this improvement might obscure any deterioration in performance in the "report" condition.

This difficulty is an inevitable consequence of the experimental design, in which subjects were used as their own controls. It cannot be overcome by reversing the order in which the "report" and the "no-report" conditions are presented, since this would introduce a further complication which the author particularly wished to avoid and which was mentioned in the Introduction. This is that, if subjects perform the "report" condition first, it becomes extremely difficult to interpret their behaviour in the "no-report" condition afterwards. Once having had to take notice of the first signal, they may go on taking notice of it, and may have to inhibit the previously learned response to it. An alternative solution might be to abandon the use of subjects as their own controls and to use each subject for one condition only. This would, of course, require a considerably larger number of subjects.

Some idea of the importance of any practice effect may be gained by comparing the results of the two successive "no-report" conditions, Sessions 2 and 3.

If one considers the difference in mean reaction times between Sessions 2 and 3 for the three subjects in the regular interval situation, for each of the five values of interval used, and if one considers, as a first approximation, the direction of the differences only, Session 3 shows 10 cases of improvement and five cases of deterioration compared with Session 2. The corresponding figures for the three subjects in the random interval situation are 11 improvements and four deteriorations. In neither case are there a sufficient number of changes in one direction to establish a significant difference at $p = 0{\cdot}05$ by the simple sign test (Siegel, 1956).

To go further and consider the magnitude of the differences, an approximate measure can be obtained by taking the difference between the mean reaction time in Session 2 and the mean reaction time in Session 3, for each subject at each interval, and using the mean of these differences as an overall measure. For the regular interval situation this gives a mean improvement of approximately 6 millisec. in Session 3 compared with Session 2, and for the random interval situation the mean improvement is approximately 10 millisec.

Thus the magnitude of any practice effect would not have been sufficient to obscure a real difference between the "no-report" and the "report" conditions, if this had existed.

CONCLUSIONS

The conclusions the author draws from these results may be summarized:—

(1) In the random interval situation there was a disturbing effect on the response to the second signal, caused by interpolating a previous signal, even though subjects had been told that the first signal was completely irrelevant to their task. The effect was significant only at the shortest interval, 50 millisec.

(2) This effect was not present when the two signals were presented with a fixed interval between them, even though the values of these intervals were the same as in the random interval situation.

(3) For both random and regular interval situations, it was possible to report on the nature of the first signal, after responding to the second, without causing any further delay in the response to the second signal.

From these conclusions one may draw the following theoretical implications:—

(1) The fact that the first signal gave rise to a delay in the random interval situation supports the notion of "central blocking" proposed in the intermittency theory. However, the duration of the block was appreciably shorter than the value of 200 millisec. proposed by the author for simple RT situations. This suggests that the blockage caused by paying attention to a signal may not be as long as when an overt response to it has to be organized.

(2) Since it is proposed that any sort of central analysis of a signal causes a block of a certain duration, then the fact that the first signal could be identified without causing a delay in the response to the second is taken by the author to imply that it is held in store *before* being analysed. Since it is unlikely that this storage was completely peripheral (e.g. in the form of after-images) the results suggest that the central decision system has some kind of short term store, in which signals are held before they are admitted for analysis, after the manner suggested by Broadbent (1958). The order in which signals arrive does not determine the order in which they will be analysed.

(3) To explain the fact that, in the regular interval situation, there were no significant delays at the shortest interval (but for one exception), one might suggest that, if the subject is set to deal with another signal at a fixed interval afterwards, he can hold the first signal in store, and postpone its analysis, until the second has been dealt with. If, however, the time at which the second signal will occur is not known by the subject beforehand, there is a tendency to start the analysis of the first signal. Hence, if a second signal does arrive shortly after the first, it will find the central mechanisms blocked.

This research was carried out while the author held a visiting fellowship at the University of Michigan from the Department of Scientific and Industrial Research. The experiment was conducted as a part of a programme of research supported by the U.S.A.F. Office of Scientific Research (Dr. Paul M. Fitts, Supervisor). Use of the facilities of the Mental Health Research Institute (Dr. J. G. Miller, Director) is also gratefully acknowledged.

REFERENCES

BROADBENT, D. E. (1958). *Perception and Communication.* London.

CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, **38**, 56–61.

CRAIK, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, **38**, 142–8.

DAVIS, R. (1956). The limit of the "psychological refractory period." *Quart. J. exp. Psychol.*, **8**, 24–38.

DAVIS, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.*, **9**, 119–29.

DAVIS, R. (1959). The role of "attention" in the psychological refractory period. *Quart. J. exp. Psychol.*, **11**, 211–20.

ELITHORN, A., and LAWRENCE, C. (1955). Central inhibition. Some refractory observations. *Quart. J. exp. Psychol.*, **7**, 116–27.

FRAISSE, P. (1957). La période réfractaire psychologique. *L'Annèe Psychol.*, **57**, 315–28.

HALLIDAY, A. M., KERR, M., and ELITHORN, A. (1960). Grouping of stimuli and apparent exceptions to the psychological refractory period. *Quart. J. exp. Psychol.*, **12**, 72–89.

HICK, W. E. (1948). Discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36–57.

MARRILL, T. (1957). The psychological refractory phase. *Brit. J. Psychol.*, **48**, 93–7.

SIEGEL, S. (1956). *Non-Parametric Statistics for the Behavioral Sciences.* New York.

WELFORD, A. T. (1952). The "psychological refractory period" and the timing of high speed performance; a review and a theory. *Brit. J. Psychol.*, **43**, 2–19.

WELFORD, A. T. (1959). Evidence of a single channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, **11**, 193–210.

THE EFFECT OF INTERMITTENT ILLUMINATION ON THE VISUAL ACUITY THRESHOLD

BY

K. GIBBINS and C. I. HOWARTH

From the Department of Psychology, University of Hull

Senders (1949) reported that with intermittent illumination short flashes needed less energy than larger ones to make a visual acuity target visible. This apparent breakdown of the reciprocal relationship between time and intensity, which has been found almost universally by earlier workers under a very wide range of experimental conditions, raised a number of interesting theoretical issues such as the relationship between visual acuity and intensity discrimination thresholds, and the plausible importance of eye movements in each. Unfortunately we have been unable to repeat Senders' result in spite of the most determined efforts to reproduce her experimental conditions. We are unable to suggest why our results differ from hers, but feel that our experiments at least show that her data are very difficult to reproduce and hence no general theoretical argument can be based on them.

In view of this we felt at liberty to predict results on the basis of a model, we had tentatively constructed, which we regarded as integrating Bloch's and Talbot's laws. This enabled us to make predictions about the effect of frequency of repetitive stimulation on thresholds. Experiments were carried out to ascertain how far these predictions were substantiated. Agreement was remarkably good.

INTRODUCTION

Senders (1949) reported that with repetitive illumination, flashes of light which were short in relation to the repetition time required less light energy than longer ones for the subject to be able to distinguish a grating. This is in opposition to what would be expected on the basis of Bloch's law. Bloch's law applies to single flashes of light and states that the energy required for threshold is constant provided the flash is shorter than a critical duration, i.e. increasing the duration of a flash allows the decreasing of intensity proportionately for the same visual effect. Well above the critical duration variation in duration is irrelevant and intensity is constant.

This law is one of the best established in the study of vision. It holds for differential and absolute thresholds, for different colours of light (Rouse), for the perception of a number of dots. It also holds for acuity targets (Niven and Brown, 1944; Graham and Cook, 1937).

Senders, used repetitive flashes for which the analogous law is Talbot's law. This states that for lights flickering *above fusion* the mean intensity over time is matched in brightness by a steady light of the same mean intensity. Thus a flashing light may consist of short bright flashes or of long proportionately dimmer flashes and yet have equal brightness. Now Talbot's law applies to matched brightness not to thresholds, but it has usually been regarded as a corollary of Bloch's law, brightness and thresholds being regarded as two indices of effective intensity.

Senders' findings bring this relationship between the two laws into question. For her evidence indicates that for a threshold criterion with repetitive stimulation both *above* and *below* fusion the expected result (*reciprocity* of duration and intensity) does not occur. Her explanation (Senders, 1954) in terms of eye movement is very plausible, but her findings still ruin the generality of the relationship.

Aim of the paper

The aims of this research are threefold.

(1) Primarily it was simply a question of trying to repeat Senders' findings.

(2) Secondly, in view of a failure to do this easily the aim was to determine what factors are responsible for her effect.

(3) As none of our experiments were able to produce this effect, the aim became simply to find out how the intensity threshold for an acuity task varied with the frequency and flash duration of repetitive illumination.

Attempts to repeat Senders' results

It did not occur to us at the very beginning of this series of experiments to doubt Senders' findings; our first experiment sought to repeat her results under daylight illumination.

EXPERIMENT 1

Apparatus. The first pilot experiment involving only one subject was carried out with very crude equipment. The adaptation illumination was simply the daylight pouring into the other end of the laboratory. The equipment utilized a piece of Perspex on which lines had been scribed roughly parallel, but of no consistent width, the shadows of which, cast by the light of a pointolite via a wedge and lens system on to a paper screen, constituted the target. The lens system brought the light to a focus where it was chopped by a variable sectored wheel, driven by a constant speed motor through a series of belt gears.

Procedure. The subject simply varied the intensity by means of the wedge until he could just see the lines on the screen, for any given size of sector (duration of flash).

Results

The results seemed consistent with Bloch's law.

EXPERIMENT 2

It was decided to repeat Senders' findings under conditions of dark adaptation (as she had) and then under comparable conditions and with the same equipment except that the eye would be light adapted. It was felt that if the pilot experiment result were valid (i.e. phenomena followed Bloch's law), this would be very interesting in view of Senders' hypothesis of eye movements as we would expect, if anything, a greater effect under light adapted conditions.

Apparatus. The equipment was, therefore, modified and generally tidied up, the light source was a 6-volt car bulb whose light passed through a lens and was brought to a focus where the same variable sectored disc could chop the light sharply. Near this point, too, was a neutral density wedge. The light then passed through a more carefully made Perspex target of carefully etched lines, casting their shadows on a small screen. The whole apparatus was carefully screened—mostly with black painted cardboard tubes. The subject was also screened by means of a wire and black paper curved screen.

Procedure. The subject was dark adapted for 20 min.

All settings were carried out by the light of a red hand lamp after the screened subject had closed his eyes.

The subject was always in *total* darkness except for the stimulus.

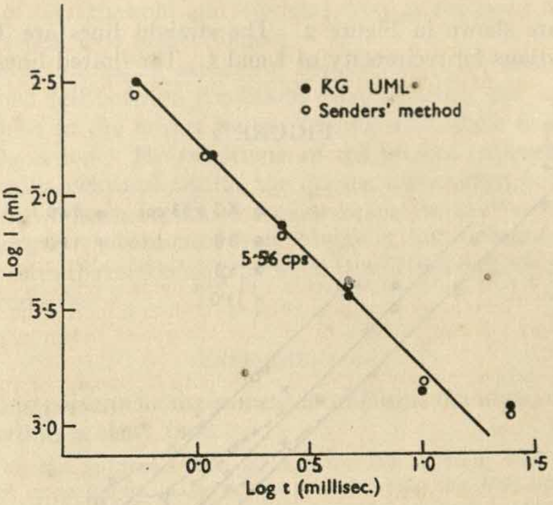
To test the effect of adaptation two procedures were used. In one, the procedure was similar to Senders'—the train of flashes was of indefinite length and the subject raised the intensity until he could clearly see the lines, waited 5 sec. and then lowered the intensity and quickly raised it to threshold level. In the other, 1 sec. bursts were given every 10 sec., to allow maximal adaptive recovery, the subject simply raised the intensity until he could just see the line.

Durations were randomized as were the conditions. Twenty estimations were made per point. Only one subject was used. Durations were measured by photocell and oscilloscope and the repetition rate checked by automatic counter.

Results

The data shown in Figure 1 are clearly consistent with Bloch's law. This result was interesting in that no evidence for Senders' findings was shown even for the dark adapted eye if only from one subject.

FIGURE 1



Results for subject Keith Gibbins.

Intensity threshold for an acuity target using repetitive flashes of various durations occurring 5·56 times per sec.

The filled circles represent results using the upward method of limits. The open circles represent the results obtained using Senders' method. Twenty estimations per point. Standard errors all smaller than 0·1 LU.

The line represents the reciprocal relation between I and *t* (Bloch's law).

EXPERIMENT 3

It was decided to make the equipment more like Senders' before carrying out more experiments.

Apparatus. The light source was again a pointolite giving a smooth output as checked by photocell and oscilloscope. A new disc was made, of aluminium alloy giving a maximum LTF of 0·43. It actually consisted of two identical discs with the same size slot cut in each, each being separately counterweighted. The LTF was varied by setting the discs at various relative angles, thus varying the size of the slot. Direct measurement and size of slot, together with the speed of the wheel, as well as calibration with photocell and scope gave us the durations of the flashes for various slot size settings.

More important a new grating was made. This consisted of a 35 mm. film of a large grating made of a sheet of white paper on which strips of black paper were pasted. The negative was mounted on a film strip viewer from which the lens had been removed and the subject viewed the negative directly. The negative was evenly illuminated by the light beam falling on the diffusing screen of the viewer. A mask cut down the apparent size of this grating to that Senders' reports as optimal—a 1° × 1° square each bar of the grating subtending 6·4 min. of arc. The target appears to have been in all respects identical with Senders'. Randomized durations were again used.

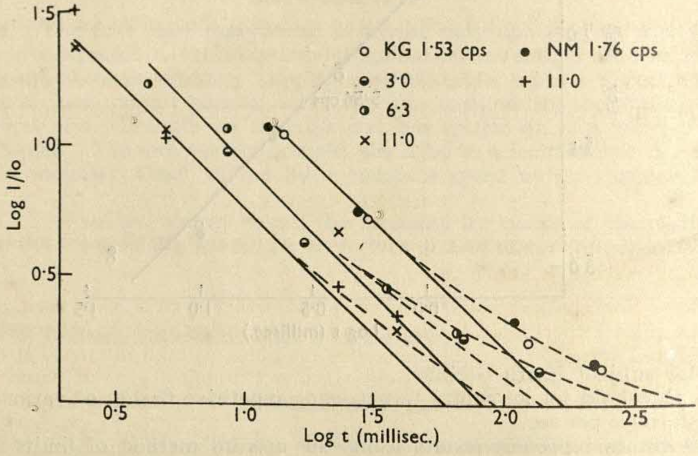
Procedure. Procedure was as before, the upward method of limits.

The possibility also arose that another feature of her procedure was responsible for her effect. This was the fact that she did not use a randomized design. Owing to the fact that with her equipment the size of the sector could not be changed and that to change the duration it was necessary to replace the disc, she did 10 trials at each duration, beginning with steady light and then going through the LTFs in descending order and then "in an attempt to balance out sequential effects, groups of 10 trials were presented in ascending order of LTF." This would mean that any practice or adaptation effects which were not linearly related to time or order would have an effect on the results and might even produce Senders' effect.

Results

The results are shown in Figure 2. The straight lines are Talbot's law lines giving the predictions for reciprocity of I and *t*. The dotted lines are for guidance only. They meet the time axis at points which represent the duration of one complete cycle and are otherwise drawn to pass as near as possible to the points for the given frequency. Only one line has been drawn for the lowest two speeds (1·53 and 1·76 c.p.s.).

FIGURE 2



Results using an equipment as much like Senders' as could be ascertained. Subjects —Keith Gibbins and Neville Moray.

Senders' method used. Twenty estimations per point.

Each curve was obtained on a different occasion.

The dotted curves meet the ordinate at duration representing one cycle. They are *not* theoretical curves.

The straight lines represent Bloch's law.

The general conclusion must be that there is no evidence for any large improvement on Talbot's or Bloch's law for results of low LTF's such as Senders reports.

As each curve was obtained on a separate occasion no reliance can be placed upon their relative positions. The expected general finding of determined threshold for a given duration as frequency increases is found, but the middle two frequencies 3·0 and 6·3 c.p.s. cannot be said to be significantly different.

Discussion

Again it is just possible that there is some difference between the present equipment and Senders' which results in our subjects having smaller eye movements and thus not allowing the operation of the process Senders thinks is responsible for her effect. The experiment was, therefore, repeated on two subjects with a grating of lines of less than half the size (2·3 min.) and also with a hairline object (32 sec.). This should have counteracted any such decrease. However, no significant deviation from Bloch's law in any direction was apparent in the results.

Here we seem to be at an impasse. Senders has reported certain findings and put forward an entirely adequate and plausible theory to explain them. In addition, she is apparently supported by two sets of independent evidence, the work of Nachmias (1958) and of Gerathewohl and Taylor (1953) as reviewed by Senders (1954) which seem like repetitions of her finding.

One of the most interesting aspects of Senders' reported findings was that the results she obtained fell between conditions where Bloch's law applies and where Talbot's law applies; in the former we are dealing with single flashes, in the latter with flashes above fusion. The conditions of the present experiments do not give Senders' theoretically awkward results, the question then must be answered "What does happen when we have repetitive flashes below fusion?"

Some indication can be obtained from Figure 2, but as these represent curves taken on different days the relationship between them is not reliable, due to day-to-day fluctuation, for example.

EXPERIMENT 4

It was decided to perform an experiment in which results from various frequencies were obtained, within a short time.

Apparatus. For the subject's comfort a projection system was used, in principle similar to the pilot experiment. The screen was let into the wall between two rooms. The dark adapted subject saw only the screen on which at high intensity stimulus levels bright bars were seen caused by the light passing through a grid of wires designed to give equal light and dark bars of visual angle 20 min. on a circular screen of 8° visual angle in diameter.

A constant speed motor drove a series of belt-driven gears which could give any of six speeds. These speeds were calibrated by photocell and oscilloscope and checked by stroboscope. Interchangeable discs of various sector sizes gave pulse-to-cycle ratios (light time fraction) of 0·1, 0·3, 0·5, 0·7 and 0·9.

The light source was a pointolite checked for ripple by photocell and oscilloscope.

The order of the discs were randomized within a set and for each disc determinations were made for three speeds in random order. Ten estimations were made per point. Three subjects were used. There were two experiments with speeds of 9·62, 4·24 and 1·67 c.p.s. and 4·81, 2·13 and 0·835 c.p.s. respectively.

The subject after dark adaptation raised the intensity until he could just see the bars, then signalled for the reading to be taken.

Results

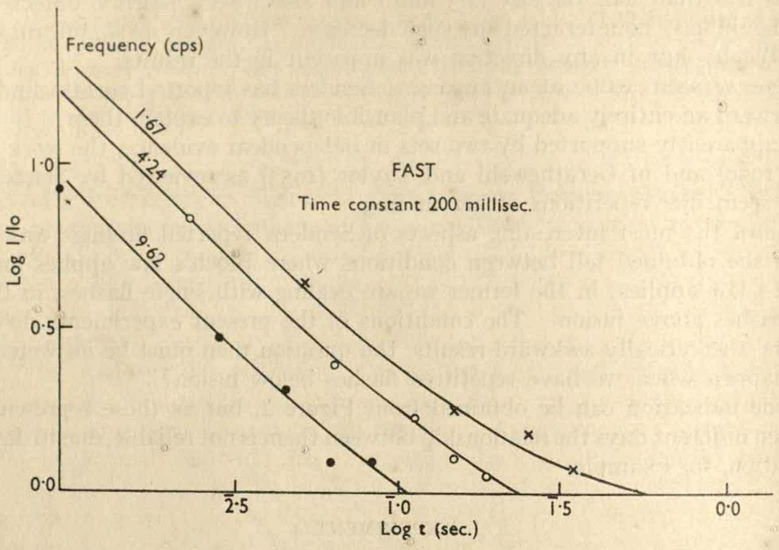
The results are shown in Figures 3 and 4.

DISCUSSION

Influenced by the present mode of considering the eye as an attenuating system the authors have developed a model of the behaviour of the visual system in the threshold situation based on the idea of a series Resistive-Capacity network (Low-pass-filter) or more generally as a system whose exponential behaviour has equal rise and decay times (time constant).

In terms of the circuit an applied intensity is seen as analogous with an applied voltage (V_a) the threshold may be represented as a critical value of effective intensity after attenuation analogous to an instantaneous voltage V_i across the capacitor;

FIGURE 3



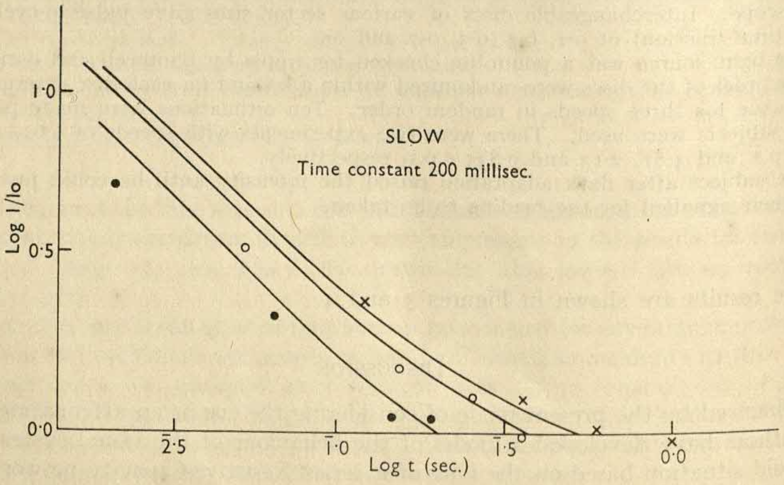
Intensity thresholds for acuity targets with intermittent illumination of 1·67 (crosses), 4·24 (open circles) and 9·62 (filled circles) c.p.s.

Mean of results of three subjects each of whom made 20 estimates per point. Senders' balancing method.

I_o was estimated.

Curves are theoretical assuming R.C. model with a time constant of 200 millisec.

FIGURE 4



Intensity thresholds for acuity targets with intermittent illumination of 0·83 (crosses), 2·12 (open circles) and 4·81 c.p.s. (filled circles).

Mean of results of three subjects, each of whom made 20 stimulations per cent. Senders' balancing method.

I_o is estimated.

Curves are theoretical, assuming R.C. circuit model with a time constant of 200 millisec.

now $V_i = V_a\,(1 - e^{-t/rc})$ for a single pulse, where *rc* is the time constant, thus the value of V_a applied for V_i just to reach a given value (V_c) is $V_a = \frac{V_c}{1 - e^{-t/rc}}$. This represents the Bloch's law situation, that is, it has the property that V_a asymptotes at $V_c t =$ constant from short applied voltage and at $V_c =$ constant for long applied voltage. Results for train of pulses will have steady-state conditions varying with the frequency.

The interest of the data of Figures 3 and 4 lies for the authors in this relationship with these theoretical curves.

It should be noted that as it stands the model is quite rigid. There are two possible variations. Variation in the time constant will simply have the effect of moving the base line by a proportional change in time constant and of changing the labelling of the curves by a like amount. Any variation in criterion voltage (V_c) will move the curves proportionately up and down the abscissa.

The curves represent the predicted theoretical result assuming a time constant of 200 millisec. To know how to interpret the relevance of the data it is necessary to consider the assumptions made. The time constant was determined by fitting the slowest frequency curve to that of the theoretical curve of a single flash, from which it would not be expected to diverge significantly. Owing to the imperfect accuracy of this procedure the time constant used was taken as a round figure (200 millisec.).

Theoretical curves were then calculated for the actual frequencies. This located the theoretical data in the time dimension.

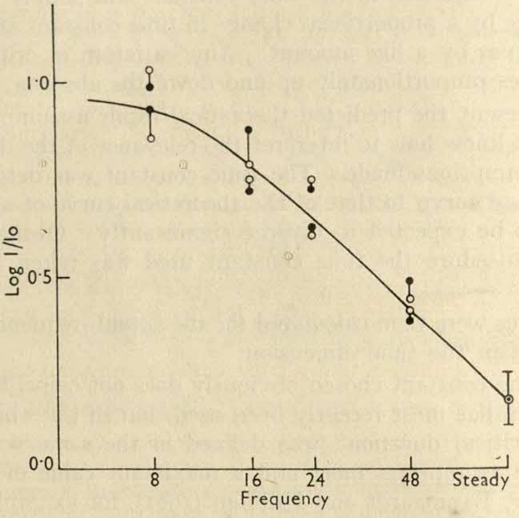
The value of time constant chosen obviously does not coincide with the critical duration as the term has most recently been used, but in the work of Graham and Margaria (1935) "critical duration" was defined in the same way as is our time constant, i.e. where asymptotes meet and a maximum value of 100 millisec. was obtained. However, Baumgardt and Hillman (1961), for example, use it to mean the duration at which actual data deviate from an $It =$ constant line. In both their data and in Barlow's (1958) this is found to be about 100 millisec. whereas the time constant varies between 100 and about 500 millisec. It should be noted that these workers used single flashes peripherally presented and determined differential thresholds. These results differ appreciably from those of Graham and Margaria (1935). Baumgardt and Hillman (1961) have, however, put forward a very plausible explanation of the discrepancy and it seems their own results are a better basis from which to work. Consequently, it seems that our choice of 200 millisec. for the time constant is a reasonable one, as it fits very well with their data.

As may be seen the points are consistent with the theoretical curve—within the limits of experimental error. The predicted more gentle change from $I =$ constant to $It =$ constant for lower frequencies, is definitely present and the data are consistent with the bunching of curves at low frequencies predicted by the theory.

This brings us back to Senders. Her results are apparently supported by the work of Nachmias (1961) who also reported that under her conditions the shorter flashes have lower energy threshold than might have been expected on the basis of Talbot's law. However, in a note (Gibbins, 1961) one of us has pointed out that his data do not require this interpretation. In fact, they are in very good agreement with the results which would be predicted on the basis of the theory presented here. He used constant duration pulses (10·4 millisec.) and varied the frequency. Some of his data plotted simply as log of relative intensity against log frequency is shown in Figure 5 along with the theoretical curve for this case. The remainder of his data have been considered in the note mentioned. The authors do not feel that Nachmias' evidence does, in fact, give any support to Senders' case, in fact, they fit our own theoretical curve rather well when allowance is made for the probable slight inaccuracy of the steady light threshold. However, discussion is continuing (Nachmias, 1961 and personal communication).

Senders has also claimed support from the work of Garathewohl and Taylor (1953). They investigated the effect of flickering light on reading matter of varying contrasts. They themselves felt their data obeyed Talbot's law (constant average intensity for equal brightness), but Senders (1954) has replotted their data, and her contention looks convincing. However, closer examination shows that this is not so. They used a card of printed matter whose contrast decreased line by line from black-on-white through black-on-grey to black-on-black. They give the *mean* reflectance of each line and Senders appears to have plotted the reflectance of the last line read against LTF of light used. However, this value gives only the intensity of the average light reflected and as is shown by the very title of the paper it is the contrast that matters. No amount of light will enable a subject to read a message if words and background are identically bright. Consequently the present authors feel the data to be irrelevant.

FIGURE 5



A replot of some of Nachmias' data.

Filled circles are the results for subject R.V., the open circles are those of subject E.G.H. In each set of results at any frequency, the upper pair of points represent results for an acuity task using short bursts of repetitive stimulation, the lower pair, a brightness matching task using a similar short train of flashes.

The curve is theoretical, using the R.C. model with a time constant of 100 millisec.

We feel the present repetition of her experiments to have been as full as could reasonably be expected, but they were completely unsuccessful in repeating her results. Some unnoted difference in experiment is probably present, but our data seems to indicate that her results are not the very important nigger-in-the-theoretical-woodpile which we first thought they were.

The theory put forward in this paper to account for Bloch's law is regarded only as tentative. The eye is a complex mechanism and so simple a model is surely wrong. However, it appears to have some promise. It is hoped in a further article to consider comparatively various theories of Bloch's law and their predictive power. We feel its relative freedom from *arbitrary* constants and its difference from other Bloch's law theories renders its use permissible in the present paper.

The work on which this paper is based was carried out under a contract with the Director General of Medical Services, Royal Air Force Flying Personnel Research Committee (Vision Sub-Committee).

REFERENCES

BARLOW, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *J. Physiol.*, **141**, 337–50.

BAUMGARDT, E., and HILLMAN, B. (1961). Duration and size as determinants of peripheral retinal response. *J. opt. Soc. Amer.*, **51**, 340–4.

GERATHEWOHL, S. J., and TAYLOR, W. J. (1953). Effect of intermittent light on the readability of printed matter under conditions of decreasing contrast. *J. exp. Psychol.*, **46**, 278–82.

GRAHAM, C. H., and COOK, C. (1937). Visual acuity as a function of intensity and exposure time. *Amer. J. Psychol.*, **49**, 654–65.

GRAHAM, C. H., and MARGARIA, R. (1935). Area and the intensity relation in the peripheral retina. *Amer. J. Physiol.*, **113**, 299–305.

GIBBINS, K. (1961). Effect of over-all duration on acuity threshold and brightness matching tasks. *J. opt. Soc. Amer.*, **51**, 457–8.

NACHMIAS, J. (1958). Brightness and visual acuity with intermittent illumination. *J. opt. Soc. Amer.*, **48**, 726–30.

NACHMIAS, J. (1961). Brightness and acuity with intermittent illumination. *J. opt. Soc. Amer.*, **51**, 805.

NIVEN, J. I., and BROWN, R. H. (1944). Visual resolution as a function of intensity and exposure time in the human fovea. *J. opt. Soc. Amer.*, **34**, 738–43.

SENDERS, V. L. (1949). Visual resolution with periodically interrupted light. *J. exp. Psychol.*, **39**, 453–65.

SENDERS, V. L. (1954). On reading printed matter with interrupted light. *J. exp. Psychol.*, **47**, 135–6.

WOODPILE REVISITED: A COMMENT ON GIBBINS AND HOWARTHS' FAILURE TO REPEAT SENDERS' FINDINGS

BY

JOHN W. SENDERS and VIRGINIA L. SENDERS*

From Bolt Beranek and Newman, Cambridge, Mass. and the Department of Psychology, University of Minnesota

Gibbins and Howarth (1962) do not get the effect that Senders (1949) did get. Gibbins (1961) disagrees with Nachmias' interpretation of Nachmias' data (1958). Gibbins and Howarth (1962) take issue with Senders' interpretation (1954) of Gerathewohl and Taylor's data (1953). The arguments about interpretation are certainly legitimate, but the failure to reproduce the effect is, at first examination, highly surprising.

That the effect is not universal is clear—Gibbins and Howarth cannot find it and Nachmias did not find it under some conditions. That it exists is also clear —Senders found it, and so did Nachmias under some other conditions. And, in Senders' experiments at least, it was a large, noticeable effect. A frequent comment of the observers, none of them naïve, was that "It seems too dim but I can still see the bars clearly." Subjectively, the grid became identifiable as such almost as soon as it became visible as a patch of light. Both the authors have clear memories of the simultaneous dimness of the target patch and clarity of the bars.

The visual mechanism is a sensitive and complicated one, reflecting subtle changes in procedure by large differences in results. The range of values of visual acuity obtained from the use of different kinds of targets, to cite but one example, bears ample witness to this fact. One dare not be "approximate" in visual experimentation; one dare not make "insignificant changes" in optics, targets, response indicators, or procedure on the assumption that these changes will have no effect. In their experiments I and II, Gibbins and Howarth used narrow dark bars on a bright field, for example, instead of equal stripes of bright and dark. (If Senders' interpretation of her data is correct, the greatest effect should probably be obtained from bright bars on a dark field.) This is a minor variation. There is a more important one. Apparently Gibbins and Howarth did not use a fixation point in any of their experiments. Nachmias and Senders used similar apparatus, and both used fixation marks, carefully controlled. At the very least, we would expect that failure to control gross eye-movements (or even head movements) would add to response variability and partially mask any effect that is present. If Senders' interpretation is correct, gross eye-movements could do more than mask the effect—they could eliminate it completely. Perhaps Gibbins and Howarth have inadvertently carried out something close to an *experimentum crucis:* by demonstrating that without fixation the effect does not occur, they have underscored the probable rôle of psychological nystagmus in enhancing the effectiveness of interrupted light.

---

* It has been a number of years since I have given much thought to theoretical problems connected with the study of visual acuity. I have therefore sought the assistance of a colleague and "close associate" who also served as one of the trial subjects in the preliminary phases of my study. His memory of the original experiment is as good as mine, and his current thinking is undoubtedly better. What follows is the product of our joint thinking. Much of the work is his, but the final responsibility remains, of course, my own. V. L. S.

Gibbins and Howarth disagree with Senders' interpretation of Gerathewohl and Taylor's data, arguing that "it is the contrast that matters." Gerathewohl and Taylor used *black* letters (rather than letters of different shades of grey) on a background of varying reflectance, so a measure of contrast and a measure of reflectance are functionally equivalent. In *all* the targets under discussion—Gibbins' and Howarth's, Nachmias', and Senders'—it is *certainly* the contrast that matters, whether that contrast is obtained by transillumination or by reflectance. The three experimental situations are in this respect identical, and theoretical considerations that apply to one should apply to all.

In summary, the effect that Nachmias and Senders found was large and reproducible. Less energy was required to see the grid than would have been used if IT = K for the acuity function. Whether the explanation proposed by Senders is correct remains to be demonstrated. Gibbins and Howarth have, by their results, suggested an experimental variable, fixation, that deserves more carefully controlled experimentation than it has yet received.

REFERENCES

GERATHEWOHL, S. J., and TAYLOR, W. J. (1953). Effect of intermittent light on the readability of printed matter under conditions of decreasing contrast. *J. opt. Soc. Amer.*, **49**, 278–82.

GIBBINS, K. (1961). Effect of over-all duration of acuity threshold and brightness matching tasks. *J. opt. Soc. Amer.*, **51**, 457–8.

GIBBINS, K., and HOWARTH, C. I. (1962). The effect of intermittent illumination on the visual acuity threshold. *Quart. J. exp. Psychol.*, **14**, 167–75.

NACHMIAS, J. (1958). Brightness and visual acuity with intermittent illumination. *J. opt. Soc. Amer.*, **48**, 726–30.

SENDERS, V. L. (1949). Visual resolution with periodically interrupted light. *J. exp. Psychol.*, **39**, 453–65.

SENDERS, V. L. (1954). On reading printed matter with interrupted light. *J. exp. Psychol.*, **47**, 135–6.

SHORTER ARTICLES AND NOTES

NOISE AND LOSS OF SLEEP

BY

D. W. J. CORCORAN

From the Applied Psychology Research Unit, Medical Research Council, Cambridge

Two experiments are reported in which human performance was compared under continuous 90 dB. white noise, after loss of sleep, under a combination of the two and under appropriate control conditions. The results suggest that noise is effective in reducing the deterioration in performance, characteristic of loss of sleep. The relevance of the result to the arousal theory of loss of sleep is discussed.

INTRODUCTION

Theories of the effects of losing sleep nearly all refer to the rôle of the brain stem reticular formation, whose function as an "arouser" or "activator" becomes depressed by periods without sleep (Williams, Lubin and Goodnow, 1959; Wilkinson, 1958). Since the system is largely dependent upon stimulation in order to function in this way, it may follow that the degree of arousal of an organism is proportional to the degree to which it is stimulated. Depth of sleep for example, is frequently measured by the loudness of sounds necessary to wake the sleeper (Kleitman, 1939).

On the basis of arousal theory one would expect that the performance of sleepy people (not only sleeping people) would be improved by high levels of stimulation. In the past stimulation has been supplied by such measures as increased incentive (Wilkinson, 1961), task complexity (Wilkinson, 1958) or punishing blasts of noise contingent upon poor performance (Kornetsky *et al.*, 1959). Each of these methods has been effective in the expected direction, but each is very much task orientated. Thus it is still questionable whether or not stimulation not associated in any way with the task would produce the same results.

High intensity noise fed continuously to the subject whilst he is performing after loss of sleep ought, on the simple arousal model, to improve his performance in relation to his level in the absence of noise. Noise at 100 dB. or more administered in this way, however, is frequently found in certain situations to lower performance under normal sleep conditions (Broadbent, 1953, 1958). The specific question then becomes, will the effects of noise summate with those of loss of sleep to induce an even greater performance decrement or will noise subtract from the performance decrement caused by loss of sleep? In order to test these expectations two experiments were performed.

EXPERIMENT I

Task

A visual version of Bakan's auditory vigilance task was constructed (Bakan, 1959). A roll of white paper 1 in. wide was passed at a uniform speed behind a slit 1 in. by ¼ in. in a black screen. Printed in succession on the roll were numbers from one to nine which gave the impression of being random. The paper roll moved at such a speed that each number was visible for one second after which it was replaced by the next number and so on. At certain places on the roll three successive odd numbers appeared which were all different, e.g. 351. These were the signals for which the subjects were instructed to watch and report verbally. The task was prolonged for 30 min. so that 1800 numbers were presented, in which there were 20 sequences of the kind described. These were interspersed irregularly in the series such that 5 occurred in each quarter of the test.

Procedure

The subjects were 31 young naval ratings aged between 18 and 27 years. All subjects were examined audiometrically before taking part in the experiment and were assigned to one of four groups at random. Each group was then tested under one of the four possible combinations of noise and loss of sleep. Group 1 was tested under fairly quiet laboratory conditions after normal sleep on the previous night. Group 2 under quiet conditions after loss of approximately 57 hr. of sleep, that is, on the third afternoon of their deprivation. Group 3 was tested whilst white noise was administered through headphones continuously through the test, and under normal sleep conditions. Group 4 was tested in noise, after loss of sleep. (Exact measurement of noise level through headphones is difficult. It has been estimated, however, that the subjects received about 90 dB. in this and the following experiment.) Headphones were not used in the quiet conditions. These conditions will be referred to as S/Q, NS/Q, S/N, and NS/N respectively.

Each subject was given the task instructions before the test and a three minute practice session in quiet was administered in order to make sure that the instructions were clearly understood.

RESULTS OF EXPERIMENT I

Figure 1 shows the average performance during each successive quarter of the test. Loss of sleep was found to have a significant effect on performance whether it was carried out under noisy or quiet conditions. A comparison of the mean scores gives a Mann–Whitney U = 6, $p = 0{\cdot}01$ for the S/Q and NS/Q comparison and a U = 9,

FIGURE 1

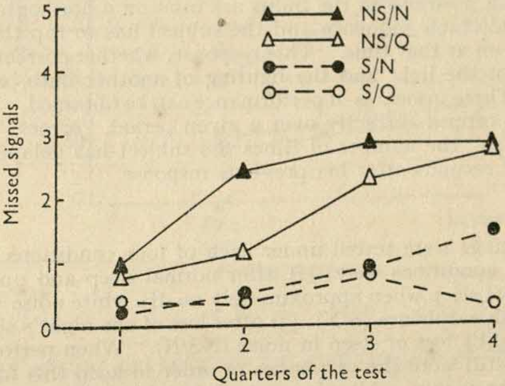
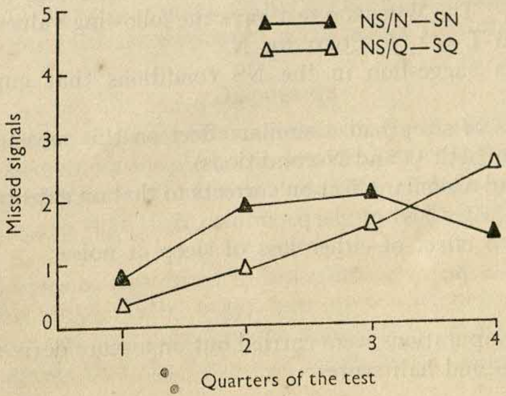


FIGURE 2



$p = 0{\cdot}02$ for the S/N and NS/N comparison. Noise, however, did not result in a significant effect on mean performance under S or NS conditions.

Loss of sleep was also found to result in a greater deterioration from the first to the second half of the test, but this occurred only under quiet conditions ($U = 4$, $p = 0{\cdot}004$). Under noisy conditions there was no significant difference in deterioration scores ($U = 23$, $p = 0{\cdot}191$). Figure 2 illustrates this result. In this figure each point on the S plot has been subtracted from the equivalent point on the NS plot for the same noise condition. Thus whilst noise was, if anything, producing poorer mean performance in the NS/N group as compared with the NS/Q group, it did appear somewhat to decelerate the decline during the NS/N run.

EXPERIMENT II

In view of the suggestion in the previous experiment that noise might have been decelerating the decline in performance after loss of sleep, a further experiment was conducted on a task previously shown to be affected by noise (Broadbent, 1953).

Task

Leonard's serial reaction task was used (Leonard, 1959). This consists of a display of five bulbs arranged pentagonally on a panel slightly inclined from the vertical. Five discs in corresponding positions to the bulbs are inset on a horizontal board. The bulbs light up in an unpredictable sequence and the subject has to tap the disc corresponding to the light which is on at that time. This response, whether correct or incorrect, results in the cancellation of the light and the lighting of another bulb, whose disc has to be tapped, and so on. Three measures of performance can be obtained. These are "corrects," the number of discs tapped correctly over a given period, "errors," the number tapped incorrectly, and "gaps" the number of times the subject has delayed his response more than one and a half seconds after his previous response.

Procedure

Twelve naval ratings were tested under each of four conditions according to a latin square design. The conditions were: (1) after normal sleep and under quiet conditions (S/Q); (2) after normal sleep when approximately 90 dB. white noise was fed continuously to the subject through earphones (S/N); (3) after loss of one night's sleep in quiet (NS/Q); and (4) after one night's loss of sleep in noise (NS/N). When performing in the Q conditions the subjects still wore the earphones in order to keep this factor constant. The task was prolonged for 30 min. All subjects had shown a satisfactory audiometric record.

RESULTS OF EXPERIMENT II

Mean scores

Gaps: (i) A greater number of gaps were scored after loss of sleep under both N and Q conditions. The Wilcoxon test gave the following values: $T = 0$, $p < 0{\cdot}01$, for Q condition and $T = 1$, $p < 0{\cdot}01$ for N.

(ii) There was a suggestion in the NS conditions that gaps were somewhat reduced by noise.

Corrects: (i) Loss of sleep had a similar effect on this measure to that on gaps ($T = 0$, $p < 0{\cdot}01$ in both Q and N conditions).

(ii) Noise also had a similar effect on corrects to that on gaps tending, if anything, to increase corrects after loss of sleep.

Errors showed no effect of either loss of sleep or noise.

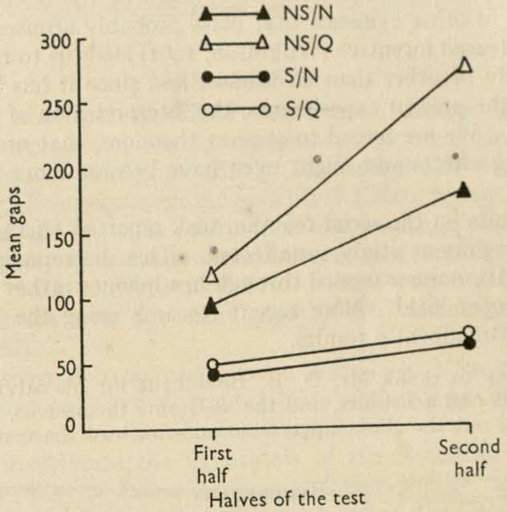
Deterioration scores

The following computations were carried out on a score derived from subtracting the first from the second half scores.

Gaps: (i) Loss of sleep showed significant effects in both Q and N conditions. The sign test gave p values of 0·006 and 0·038 respectively

(ii) Whilst there was no greater deterioration in noise than in quiet in the S condition (nine subjects in fact showing greater deterioration in noise), the NS condition showed significantly less deterioration in noise. Two of the 12 only showed greater deterioration which gives a $p = 0{\cdot}038$. This clearly indicates that noise, in the loss of sleep condition decelerated the decline in performance evident in performance after loss of sleep in quiet conditions (Fig. 3).

FIGURE 3



Corrects showed similar but insignificant confirmation of the gap results; errors showed no differences at all.

Second half scores

Gap scores attained during the second half of the test failed to show a significant difference in the NS/N, NS/Q comparison, but correct taps were significantly fewer in NS/Q than NS/N ($T = 14$, $p < 0{\cdot}05$). This again indicates that noise administered after loss of sleep lessened the decline in performance towards the end of the test.

DISCUSSION

The effect of noise in these experiments was clearly to reduce the decline during the test which is normally associated with loss of sleep. This result is in accordance with a simple arousal theory, which would maintain that noise is arousing and loss of sleep de-arousing, so that their combination results in levels of arousal between these extremes.

Both noise and loss of sleep tend to have their greatest effects towards the end of tests (Broadbent, 1958; Tyler, 1947), and, quite reasonably, it was towards the end in these experiments that noise became beneficial to sleep deprived performance. This naturally suggests that the decline in performance in noise is associated with hyper-arousal and therefore that arousal increases during the test under noisy conditions. However, there is evidence to suggest that the output of the ear declines when continuously stimulated (Hood, 1950), indicating that the intensity of sound reaching the more central areas of the nervous system declines in time. If this is the case noise ought to be more arousing at the beginning rather than at the end of a run. Indeed, a perfectly reasonable argument would be put forward identifying continuous loud noise with gradually lowered arousal. Noise might prove arousing at first, but by cutting out other stimulation, result in a progressive de-arousal as time progresses. This argument would necessitate the identification of loss of sleep and hyper-arousal. As Broadbent (1962) has pointed out, the fact that a negative interaction between noise and loss of sleep has been demonstrated suggests that whichever of the two is the arouser the other must be the de-arouser, and he concludes on the basis of other evidence that noise probably arouses. A consideration of the effects of increased incentive (Wilkinson, 1961) leads us to the same conclusion. Incentive can hardly be other than an arouser and since it has very similar effects to that of noise in the present experiments, the interpretation of noise as an arouser is clearly favoured. We are forced to suggest therefore, that noise of this kind did not lose its arousing effects and might even have become more potent in this way as time progressed.

Broadbent's results on the serial reaction task reported effects of noise on errors which were, in the present study, unaffected. This discrepancy is probably due to the fact that 90 dB. noise was used through headphones rather than 100 dB. noise projected into the open field. More recent research using the latter method has, in fact, confirmed Broadbent's results.

The author wishes to thank Mr. D. E. Broadbent for his advice, the R.N.P.R.C. for supplying subjects and assistants, and the assistants themselves, in particular C.P.O. Collecot, for carrying out the close supervision and much of the testing.

REFERENCES

BAKAN, P. (1959). Extraversion—Introversion and improvement in an auditory vigilance task. *Brit. J. Psychol.*, **50**, 325–32.

BROADBENT, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, **44**, 295–303.

BROADBENT, D. E. (1958). *Perception and Communication.* London.

BROADBENT, D. E. (1962). Possibilities and difficulties in the concept of arousal. *Santa Barbara Symposium on Vigilance*, (1961), *U.S. Office of Navy Research.* In press.

HOOD, J. D. (1950). Studies in auditory fatigue and adaptation. *Acta Otolaryngologica Supplementum XCII.*

KLEITMAN, N. (1939). *Sleep and Wakefulness.* Chicago.

KORNETSKY, C., MIRSKY, A. F., KESSLER, E. F., and DORFF, J. E. (1959). The effects of dextro-amphetamine on behavioural deficits produced by sleep loss on humans. *J. Pharmacol.*, **127**, 45–50.

LEONARD, J. A. (1959). Five choice serial reaction apparatus. *M.R.C. Report A.P.U.* 326/59.

TYLER, D. B. (1947). Effects of amphetamine sulphate and some barbituates on fatigue produced by prolonged wakefulness. *Amer. J. Physiol.*, **150**, 253–62.

WILKINSON, R. T. (1958). *Effects of Lack of Sleep on Perception and Skill.* Ph.D. Thesis, University of Cambridge.

WILKINSON, R. T. (1961). Interactions of lack of sleep with knowledge of results, repeated testing and individual differences. *J. exp. Psychol.*, **62**, 263–71.

WILLIAMS, H. L., LUBIN, T., and GOODNOW, J. L. (1959). Impaired performance with acute sleep loss. *Psychol. Monogr.*, **73**, 1–26.

THE EEG, EYE-MOVEMENTS AND DREAMS OF THE BLIND

BY

R. J. BERGER, P. OLLEY and I. OSWALD

From the Department of Psychological Medicine, University of Edinburgh

It was found that characteristic frontal EEG waves significantly often precede the rapid eye movements of dreaming. The rapid eye movements were absent during dream periods of three men with life-long blindness, and of two men, 30 and 40 years blind respectively, but were present during dream periods in three men blind only 3, 10 and 15 years respectively.

INTRODUCTION

In recent years, numerous American and French workers have described the invariable cyclical appearance, during continuous nocturnal sleep, of light sleep patterns in the electroencephalogram (EEG), accompanied by frequent and characteristic clusters of rapid conjugate eye movements (REMs) and dreaming (e.g. Dement and Kleitman, 1957*a* and *b*). The REMs have been claimed to represent scanning movements with respect to the dream visual imagery (Dement and Kleitman, 1957*b*). The "REM periods" with accompanying EEG picture vary in duration (20 min. is common) and four to six occur per night, together with characteristic muscle tonus changes (Berger, 1961).

Schwartz and Fischgold (1960) and Jouvet, Michel and Mounier (1960) described brief appearances of 2 to 3 c./sec. frontal EEG activity, having a characteristic "saw-toothed" wave form associated with the REMS.

We decided to investigate the hypothesis of the American workers that the REMs represent what may be called "looking-at-dream-picture" movements. First, by a closer examination of the time relations between the "saw-toothed" frontal EEG activity and the onset of individual REM clusters; secondly, by studies of the EEG and eye movements during nocturnal sleep in the blind.

Jastrow (quoted by Ramsay, 1953) studied the dream reports of persons who had become blind at varying periods after birth. Those who became blind later than early childhood maintained visual imagery but there was a progressive decline of reports of visual imagery as the individual became older. A few individuals only were said still to report some visual imagery after 40 years of blindness.

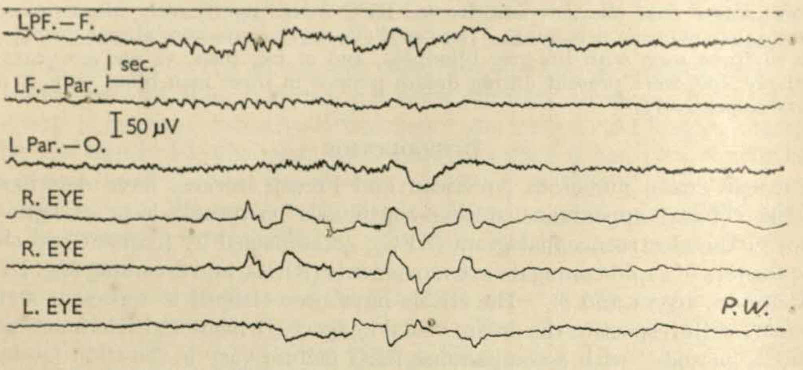
METHOD AND RESULTS

The EEG "saw-toothed" waves

In the course of other studies of all-night sleep we observed that among 36 normal persons, on whom suitably located electrodes had been used, were 23 who showed the 3 c./sec. frontal "saw-toothed" waves very prominently. It seemed clear to us that generally these brief bursts of EEG waves *preceded* individual REM clusters. In order to confirm this hypothesis, we recalled the first of our subjects from whom we had recorded the saw-toothed waves, and recorded one further entire night's sleep. The EEG was recorded using a left antero-posterior chain of silver cup electrodes, and eye movements by means of the most anterior EEG electrode and two other electrodes fixed just below the outer canthus of each eye. The diagonal and asymmetrical electrode placement had the advantage of economy, and allowed both horizontal and vertical REMs to be recorded on the same channels. Figure 1 illustrates.

Subsequently, with the eye movement channels and then the EEG channels of the record alternately obscured, two of us (I.O. and R.J.B. respectively) independently marked the record where it was judged that each "saw-toothed" burst and each REM cluster began and ended. Figure 2 indicates that our hypothesis was confirmed; the onset of "saw-toothed" frontal EEG waves is significantly ($\chi^2 = 229{\cdot}4$, $n = 2$, $p < 0{\cdot}001$) related to the five seconds preceding a REM cluster. This result was obtained by comparing the observed frequencies with those which would be expected to occur by chance, the latter being proportional to the total duration of each condition.

FIGURE 1



An example of "saw-toothed" frontal activity preceding the appearance of a brief REM cluster by 3 sec.

LPF—F = left pre-frontal—frontal.
LF—Par. = left fronto—parietal.
L Par—O = left parieto—occipital.

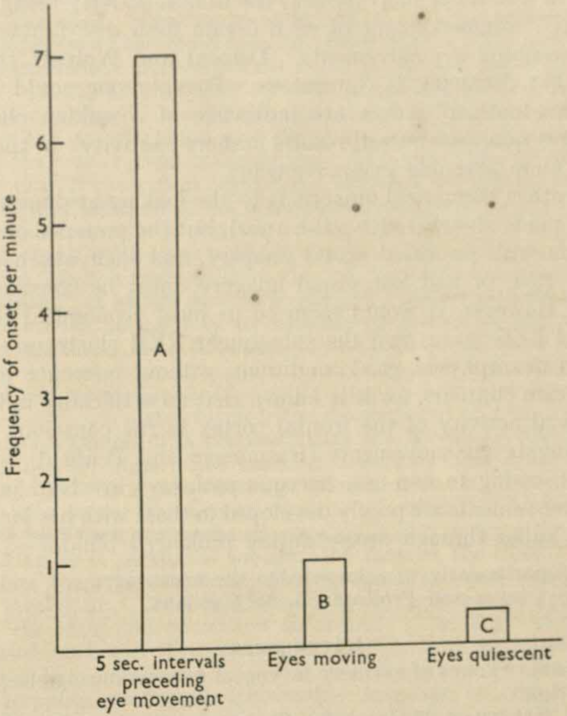
Blind subjects

An antero-posterior chain of silver cup electrodes was used to record the EEG. A pair of electrodes was fixed at each outer canthus, to give four different electrode combinations, to enable us to recognize the occasional electrode artefact (these sometimes look like REM potentials). Eight blind male adults were studied, each having two entire nights of recording.

Three men had had life-long blindness; two, aged 25 and 30, on account of severe corneal scarring (probably resulting from neonatal gonorrhoea) and one aged 17, on account of bilateral optic atrophy since the age of 18 months or earlier. These men had no conception of visual imagery. A 60-year-old man had been blind for 30 years owing to corneal scarring, and a 52-year-old man for 40 years owing to traumatic retinal detachment and cataracts. The latter two men both stated that they had lost the ability to picture things and claimed that no visual imagery was present in their dreams. Each was first shown to produce clear electro-oculographic potentials when the eyeballs were moved voluntarily. None of these six men had REMs in their cyclically recurring light sleep periods, during which, to judge by their subsequent reports, they nevertheless dreamed of non-visual fantasy experiences.

Three men aged 23, 33 and 40, totally blind for 10, 3 and 15 years respectively claimed still to experience visual imagery. Each had REMs in their cyclically recurring light sleep periods differing in no way from those of normal sighted people, and each when wakened during REM periods recalled visual dreams. As an illustration, the man who had been blind for 15 years (who had particularly active REMs) described how he dreamed he was in the swimming bath, how he did the "crawl" rapidly to the other end where he could see his friend's wife ("her white skin and black costume, it was all very vivid") and splashed water teasingly over her. Of the eight blind men, only this last man had "saw-toothed" waves.

FIGURE 2



Frequencies of occurrence per minute of the "saw-toothed" activity during 5-sec. intervals preceding REM clusters, during REM clusters, and when the eyes are quiescent during the cyclically recurring light sleep periods.

A: 28 instances of onset of saw-toothed activity in a total of 4 min. (namely, 5-sec. periods preceding a total of 48 REM clusters).

B: 20 instances of onset of saw-toothed activity during the course of the 48 REM clusters of total duration 20 min.

C: 21 instances of onset of saw-toothed activity during a total of 72 min. while the eyes were quiescent during REM periods.

Discussion

Wolpert (1960) implies a theory of peripheral control of dreaming in the statement: "Thus REMs appear to be specific to the manifest dream content, thereby confirming the suggestion put forward by Ladd as early as 1892 that during dreaming the eyeballs move gently in their sockets, taking various positions induced by the retinal phantasms as they control the dream."

A peripheral theory of this nature does not commend itself to us and we find it difficult to reconcile our observations concerning the "saw-toothed" EEG waves with a looking-at-the-dream-pictures hypothesis. It could perhaps be argued that dreaming occurs in a series of short flashes: the frontal activity being a physiological concomitant of the commencement of each dream flash, the fantasied contents of which provoke scanning eye-movements. Dement and Wolpert (1958), however, argue strongly that dreaming is continuous. Possibly one could argue that the frontal EEG "saw-toothed" waves are indicative of a sudden change of neurophysiological status quo, which itself results in more "activity" in the dream events, so calling forth more scanning eye-movements.

A number of other theoretical objections to the looking-at-dream-pictures hypothesis have been made elsewhere (Oswald, 1962), but the presence of REMs in those blind subjects who still possessed visual imagery, and their absence in those who had either never had, or had lost visual imagery, must be considered to support the hypothesis. However, it would seem to us most economical to suppose that both the burst of EEG waves and the subsequent REM cluster are indicative of a sudden change of neurophysiological conditions, without reference to awareness of, or response to dream contents, for it is known that an artificially induced change of electro-physiological activity of the frontal cortex in the conscious human can be followed by conjugate eye-movements (Rasmussen and Penfield, 1948). It could be supposed that, owing to non-use, nervous pathways involved in the execution of conjugate eye-movements are poorly developed in those with life-long blindness, or, once established, suffer through disuse during prolonged blindness.

We should like particularly to acknowledge the encouragement and facilities given us by Professor J. Drever and Professor G. M. Carstairs.

References

Berger, R. J. (1961). Tonus of extrinsic laryngeal muscles during sleep and dreaming. *Science*, **134**, 840.

Dement, W., and Kleitman, N. (1957*a*). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroenceph. clin. Neurophysiol.*, **9**, 673–90.

Dement, W., and Kleitman, N. (1957*b*). The relation of eye movements during sleep to dream activity, an objective method for the study of dreaming. *J. exp. Psychol.*, **53**, 339–46.

Dement, W., and Wolpert, E. A. (1958). The relation of eye movements, body motility and external stimuli to dream content. *J. exp. Psychol.*, **55**, 543–53.

Jouvet, M., Michel, F., and Mounier, D. (1960). Analyse électro-encéphalographique comparée du sommeil physiologique chez le chat et l'homme. *Rev. neurol.*, **103**, 189–205.

Oswald, I. (1962). *Sleeping and Waking: Physiology and Psychology*. Amsterdam.

Ramsey, G. V. (1953). Studies of dreaming. *Psychol. Bull.*, **50**, 432–55.

Rasmussen, T., and Penfield, W. (1948). Movement of head and eyes from stimulation of the human frontal cortex. *Proc. Ass. Res. nerv. ment. Dis.*, **27**, 346–61.

Schwartz, B. A., and Fischgold, H. (1960). Introduction à l'étude polygraphique du sommeil de nuit (mouvements oculaires et cycles de sommeil). *Vie méd.*, **41**, 39–46.

Wolpert, E. A. (1960). Studies in psychophysiology of dreams. *A.M.A. Arch. gen. Psychiat.*, **2**, 231–41.

BOOK REVIEWS

Plans and the Structure of Behaviour. By George A. Miller, Eugene Galanter and Karl H. Pribram. New York. Henry Holt. 1960. Pp. x + 226. $5.

A year spent in discussion at the Centre for Advanced Study in the Behavioural Sciences at Stanford formed the basis of the arguments presented here. The book contains a review, at a very general level, of certain similarities between behaviour and the abilities of modern computers and of the implications of these for the explanation of behaviour.

The argument is, throughout, pitched at a very general level indeed. A distinction is made between Images and Plans, the former term being drawn from the work of an economist, K. E. Boulding. The Image is defined as "all the accumulated, organized knowledge that the organism has about itself and the world." Plans are "any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed." As the authors point out, the Image corresponds roughly with whatever mechanism is concerned, in cognitive theories of behaviour, with the storage of information about the environment; and Plans are closely analogous to programmes designed to direct the deliberations of computers. It is clear that the authors have little faith in the future of S–R theories in psychology, although the reasons for this scepticism are, at best, only briefly mentioned. A preference is expressed for the more "cognitive" kinds of theory, again rather as a premise for the subsequent argument. It is pointed out, however, that cognitive theories tend to contain an executive defect. The cognitive animal may be wise, erudite and highly motivated, but it is also prone to be the most ineffective of creatures. It knows what it wants and even perhaps what it should do to get it, but it suffers from an inability to translate its intentions into action. There is no doubt that this difficulty is to be found, in one form or another, in most cognitive theories and that it constitutes an important defect which has often passed without remark. The problem is stated in the book with unusual clarity and the analysis therefore gets off to a good start. As a step towards the development of a theory which does not suffer from this disadvantage, it is proposed that the nervous system consists in part of a hierarchy of "TOTE" units. The function of each of these is to compare an input with some standard, perform a prescribed operation to reduce a discrepancy, and, should this be successful, pass on the transformed input to a further TOTE unit. Thus, a series of relatively simple feed-back processes perform a series of transformations on a particular input. The selection of a particular sequence of these is the "plan," or programme, and is the means by which information, both perceptual and memorized, is transformed into appropriate behaviour. Level in the hierachy of these units corresponds to degree of abstraction of the tests and operations concerned. The larger part of the book is devoted to the development of these ideas in connection with goal-directed and "intentional" behaviour, values, instinctive behaviour, the acquisition of skills, memory, certain pathological and hypnotic phenomena, personality, language, concept formation and the use of logical skills such as those required in chess. There is a final chapter concerned with speculative physiological embodiments for these mechanisms.

It is very difficult to know what to make of all this. The argument is always stimulating and it often seems that old problems are being re-phrased in an important and interesting way. This impression is strengthened by the knowledge that some of the plans discussed have actually been constructed in detail and made to work in computers —e.g. in the case of chess. It is easy to mistake programme for achievement in this kind of work and, although the authors in several places specifically warn the reader that programmes of the degree of complexity required are yet to be written, it remains necessary constantly to bear in mind that, for the most part, the background logical spadework has not even been attempted, let alone completed. In the sense in which the words are used, it is a logical truth that behaviour is controlled by plans; this logical truth becomes of psychological interest, as opposed to being of interest to psychologists, when these plans are worked out. This, of course, the book does not claim to do in any detail. In these circumstances it is necessary to scrutinize the application of the ideas rather closely. To what extent is this something more than yet another technical re-description of behaviour?

The empirical context of the argument may be assessed on two grounds: (*a*) on that of predictive power, and (*b*) by contrast with other theoretical works. It is not altogether easy to be certain, in either of these ways, that a significant contribution has been made.

The authors are themselves aware that it might be objected that their account is vacuous because no particular predictions follow from it. They include this objection with two others—namely, those of complexity and lack of a certain kind of objectivity, in a chapter on the ways in which plans or programmes are developed. The objections of complexity and lack of objectivity, which appear formidable, are met with ingenious and plausible arguments. But relatively little is said in reply to the third objection, for only the limited claim is made that certain kinds of experimental technique might reveal the particular plans according to which a given subject was operating. That is, it is not suggested that the system described in the book is open to experimental test, but only that, given this logical framework, the parameters might be ascertained by empirical study. A similar conclusion may be drawn from brief references to such theories as those of Hebb and Hull. Objection is made to these not because of a failure of correspondence with the facts; it is argued instead that modern knowledge of the working of complex machines shows that such psychological theories are couched in the wrong terms, are mechanistically too simple, to provide an adequate account of the central processes of the nervous system. This may be so. But if the present account and these alternative theoretical formulations are not to be distinguished on empirical grounds, at least it must be made clear in what ways they differ as logical frameworks. This is often not obvious from the text. The question of how far the reflex analysis and that outlined in the book are logically inter-translatable is scarcely raised. In this connection, it is perhaps a little surprising that the discussion of cognitive theories is more or less restricted to that of Tolman and that more detailed attempts of the same kind, such as Deutsch's work, are not discussed. For many, though by no means all, of the difficulties with which the book is concerned might appear already to have been overcome by this more recent formulation, and, moreover, at a more concrete and predictive level than that at which they are examined here.

Thus, the book is mainly an exercise in the use of a recently developed language in which to discuss psychological problems. It is not itself a theoretical account and it is not clear how far this language permits an approach which really differs from those derived from other frameworks. In this respect it shows something of the quality of the more theoretical Gestalt work, or, say, Bartlett's book on memory. The scope and general interest of the work also place it in this class and, as usual, the argument becomes less satisfactory and more superficial as the authors venture into such less well-defined realms as those of personality and social psychology. In common with such works, it deserves attention and may be expected to have some influence on the development of the subject; yet critical reading will tend to reduce the apparent depth and originality of many of the arguments.

A. J. WATSON.

Behaviour. By D. E. Broadbent. London. Eyre & Spottiswoode. 1961. Pp. 216. 21s.

Behaviourism and psychoanalysis are near contemporaries, or at least they became widely known at about the same time, yet, whereas the concepts and methods of psychoanalysis form part of the accepted background knowledge of any well-informed person, behaviourism, particularly in its modern and more sophisticated forms, is little understood and widely disregarded. There are various reasons for this state of affairs; Watson's original formulation of behaviourism was painfully naïve; behaviourist findings have, until fairly recently, been more reliable but much less dramatic than those of the psychoanalysts; not the least important reason is that, while psychoanalysis has been well served by propagandists and popularizers, the experimental study of behaviour has had little such attention. One result of this is that all too often, even in academic circles, psychologists are assumed to be the same as psychiatrists. Mr. Broadbent's book could do much to change this state of affairs. His book is not a general introduction to experimental psychology as a whole, rather it is a study of one particular approach to psychological problems—the behaviourist one; of its origins and justification and of its development and findings. For students of the subject the book may seem straightforward, but Mr. Broadbent's treatment is always thoughtful and may often be illuminating even to the professional. For the interested layman, at whom the book is aimed, it will not be easy reading; Mr. Broadbent does not avoid difficulties or cut corners where it would be misleading to do so, yet there is no other book which could give such a reader a better picture of the aims and achievements of this sort of psychology.

In one sense the title of the book is misleading since by no means all aspects of behaviour receive attention; instinctive behaviour for instance gets little mention; this, however, is a direct consequence of Mr. Broadbent's main aim which is to concentrate on the Law of Effect. In the introductory chapter we are shown how behaviourism arose from a dissatisfaction with the methods of classical introspectionist psychology: the philosophical and practical reasons for the rejection of introspection and the adoption of behaviourism are very well outlined, giving a pretty complete answer to those who wonder why psychologists, who should be studying the mind of man spend so much time worrying about the behaviour of animals. Throughout most of the rest of the book the emphasis is heavily on the Law of Effect; we follow it from Thorndike's original formulation, through Hull's systematizing, to the introduction of internal ("mental") responses and the more tentative and sophisticated formulations now in favour. Mr. Broadbent marshalls with expertise the mass of evidence needed to develop this theme and brings his facts to bear with precision. A good deal of ground is covered, from Lashley's psychophysiology to Skinner's operant conditioning, but the great merit of the book is that by concentrating on one theme it confines the field to be dealt with, and is thus able to treat the central theme with greater thoroughness and precision than is normally possible in a book of this sort. Thus relatively abstruse subjects are discussed such as the evidence for Mowrer's two-factor theory of avoidance learning or the development of Hull and Spence's incentive factor. On the whole, these matters are clearly explained though once or twice the arguments become rather tough going; I wonder how many people with no previous knowledge could follow the passage on the fractional anticipatory goal response and Deutsch's critique of Hull's use of it?

A second advantage of Mr. Broadbent's concentration on a single theme is that he thus gives a clear picture of the development of psychological thought, or of scientific thought in general; by the setting up, testing and adoption or rejection of hypotheses. This aspect of the book is emphasized by a chapter in which Hull's and Skinner's theoretical approaches are contrasted and by the final chapter in which various theoretical approaches now popular are discussed; Mr. Broadbent shows a pleasant distaste for dogmatism about the way experimental psychology should be done.

Throughout the book Mr. Broadbent gives many common-sense examples of the application of of psychological theory which are, on the whole, very helpful. On the other hand, he often shows how experimental results from animals might be extended to apply to human beings and even, as in the case of punishment learning, motivate social reform. In itself this is interesting and appropriate to a book of this sort, but it is unfortunate that Mr. Broadbent does not make clear to the lay reader just how speculative some of these conclusions are. For instance, from the results of irrelevant incentive learning, he draws the conclusion that "one must remain rather sceptical therefore about the possibility of a physical need, such as sex, lending vigour to the manipulation of general conceptual experience"; now this is an interesting idea but clearly the experimental results are totally inadequate to support it; however the reader may assume from this sort of thing that psychologists are prone to make wild generalizations about human behaviour on the basis of animal experiments. This would be a pity, for in general the book deserves to be widely read.

M. S. HALLIDAY.

Sensory Deprivation. Edited by Philip Solomon, M.D. and others. Preface by Stanley Cobb, M.D. Harvard University Press. London: Oxford University Press. 1961. Pp. xv + 262. 46s. net.

The Manipulation of Human Behavior. Edited by Albert D. Biderman and Herbert Zimmer. London and New York. John Wiley & Sons Ltd. 1961. Pp. xii + 323. 64s. net.

The first of these books is based on a Symposium held at the Harvard Medical School in June, 1958. In their Introduction, the Editors trace briefly the growth of interest in sensory deprivation and rightly give pride of place to Dr. D. O. Hebb, who himself provides a short introduction to the first paper, by Dr. Woodburn Heron. This is followed by four further papers reporting experimental studies and four with a more clinical slant. There is a short chapter on sensory deprivation in aviation. Theory and implications are considered in a final section. Among the better known contributors are Dr. Milton Greenblatt, Dr. D. B. Lindsley, Dr. Austin Riesen, Dr. Jerome Bruner and Dr. L. S. Kubie.

When the Montreal studies of sensory deprivation first appeared, it seemed to many that we had to do with a fairly consistent effect of reduced environmental stimulation upon a variety of psychological processes. Individual differences were of course marked but were not then regarded as of primary importance. Now, ten years later, it is plain that the effects—both during and post-isolation—are a good deal less consistent than had been supposed. For instance, some authors report a high incidence and others a low incidence of significant perceptual changes. It also seems clear that the social and emotional aspects of the isolation experience are of much greater importance than was at first realized—which no doubt accounts for the interest taken by some psychiatrists. Obviously, then, what is called "reduction of environmental stimulation" involves changes at a number of levels and it is most difficult to say what can, and what cannot, be attributed to deprivation *per se*.

The present book provides a useful review of experimental work on deprivation and includes much material not previously published. On the theoretical side, however, it is a little disappointing. The three authors who contribute to the final section do little more than assimilate the broad findings to their pre-existing schemes of psychological explanation. Lindsley, for instance, suggests that deprivation may well alter the input-output level in the C.N.S., most likely *via* the reticular formation. Bruner sees it as a process which "disrupts the vital evaluation process by which one constantly monitors . . . the strategies one has learned to employ in dealing with the environment." Kubie, who has long striven to integrate psychoanalysis with neurology, thinks that deprivation may alter the distribution of afferent input as between conscious, preconscious and unconscious systems. At the present stage of things, it would seem, we can have as many theories of deprivation as we have disciplines.

The second book is stated to be the outcome of a project undertaken by the Bureau of Social Research, Inc. The preparation, we are told, ". . . was supported in part by the United States Air Force under Contract 18(600)–1797 monitored by the Rome Air Development Center of the Air Research and Development Command under Contract AF49(638)–187 monitored by the Air Force Office of Scientific Research of the Air Research and Development Command." This, clearly, is the modern world and as might be anticipated the topic is in reality brain washing (here called interrogation). There is some quite well presented material on altered physiological states, drugs and hypnosis in relation to the extraction of information. The final chapter bears the splendid title: "Countermanipulation through Malingering."

O. L. ZANGWILL.

Statistical Reasoning in Psychology. An Introduction and Guide. By R. S. Rodger. University Tutorial Press. 1961. Pp. vii + 204. 15s.

Is measurement possible in psychology? Do psychological experiments produce the kinds of data which may legitimately be treated by statistical methods? If so, what are the assumptions underlying the techniques used? What is the meaning of "statistical significance" and when should "one-tailed" and "two-tailed" tests of significance be made? Attempts to answer these and similar questions form a large and growing body of literature which may be found in various books and journals, but in few cases are clear-cut, universally agreed answers given. The time needed to read and assess the arguments presented from different points of view would consume an undue proportion of that available to the student and might well leave him with the impression that statistics should not be employed in biological work. From the practical point of view this would be unfruitful and unprofitable since biological statistics are here and here to stay. It does, therefore, behove the student and any one else using statistical methods to try to understand what he is doing when he applies statistical tests and why he is doing it.

This little book may safely be recommended to any one who is contemplating the analysis of experimental data by statistical means. It might also be salutary reading for many who are already deeply dyed with statistical manipulation!

E. G. CHAMBERS.

Rabbit Brain Research, Volume 1: Atlas for Stereotaxic Brain Research on the Conscious Rabbit. By M. Monnier and H. Gangloff. Amsterdam. Elsevier. 1961. Pp. vii + 76. £3 3s.

This book contains a stereotaxic atlas of the rabbit brain for use with an electrode system devised by the authors. For this particular purpose the atlas appears to be quite adequate; for more general use, however, the atlas will prove much less adequate since the plates illustrate sections of the brain cut at 3-mm. intervals. This may not prove too much of a drawback, since more detailed atlases are readily available in the journals. Perhaps the justification for publishing this book lies in the opportunity it affords the authors of describing their techniques for studying the brain of the unanaesthetized rabbit. Good as this account is, it suffers from certain defects. There are occasions when unnecessarily lengthy descriptions are given of simple procedures (e.g. instructions for reading co-ordinates of the caudate nucleus), and too little detail of more complex procedures (e.g. the method of making and insulating electrodes). Nor, in the latter example, is the inadequacy of the description remedied by a footnote informing the reader that electrodes may be obtained from the Department of one of the authors.

The atlas occupies 21 pages of the book and the section on techniques, 36 pages. The latter does not contain as much information as one might suppose, since the entire text is written in English, French and German. This will make the book a valuable acquisition to someone who is learning one or other of these languages, but it also increases its cost and adds nothing to its technical value. The production of this book is up to the usual excellent standards of the publishers. GABRIEL HORN.

Experimental Psychology. By B. G. Andreas. London and New York. Wiley. 1960. Pp. xi + 595. 56s.

This text is specifically addressed to students and teachers of experimental psychology and is well worthy of their notice. The first part deals with general issues of method (design and conduct of experiments, measurement and statistics, psychological scaling) and the second with selected areas of investigation (chiefly visual and auditory perception, association and learning, problem solving). The book thus differs from most of its kind in not trying to cover too much and in placing greater stress on method than is perhaps usual. Further, Dr. Andreas makes an interesting attempt to relate basic instruction in psychology to training in research methods—including even a section on the writing of research reports! Although the book has many and obvious merits, imagination is unfortunately not one of them. Until the interest of the student is caught, it is doubtful whether he can be prevailed upon to take methodological issues very seriously. Perhaps, therefore, this book will prove of greater use to those who are beginning research than to undergraduates, many of whom will find it somewhat destructive of the idea (or illusion) that psychology is an interesting subject. Yet this book is a real text, hard and thorough, and mercifully free from the prevailing belief that undergraduates studying psychology are best taught by exposure to glossy photographs of rats with implanted electrodes or young ladies in white coats administering the Rorschach test. O. L. ZANGWILL.

Learning Theory and the Therapy of Neurosis. By Ralph Metzner. British Journal of Psychology Monograph supplement No. 32. London. Cambridge University Press, 1961. Pp. 29. 7s. 6d.

As the author of this monograph points out the interpretation of psychopathological observations in terms of learning theory is not particularly novel. The use of adversive stimuli was daily practice when devils needed to be exorcised but even the modern rationale derived from animal experiments is more than 30 years old.

Nevertheless few psychotherapists use learning theory concepts in planning or describing therapeutic programmes, while those who use conditioning methods emphasise the differences, rather than the similarities, between this type of treatment and more standard methods. Does Dr. Metzner's monograph help to lessen the gap? The answer is surely yes. Metzner does not claim to be comprehensive, but his review is systematic and he does clarify many important issues. His introductory discussion of certain concepts central to learning theory is both compact and extremely lucid. Perhaps the most surprising omission in this section is the absence of any reference to imprinting, even if imprinting is regarded, not as a distinct type of learning, but as a special case which fits the general rules, it still merits a separate mention. Firstly in relation to learnt sexual behaviour; secondly in relation to psychoanalytically derived concepts of emotional development; and thirdly, and perhaps with greater relevance in relation to one treatment recovery.

Perhaps the main defect of this very useful monograph is a function of its virtues. Thus the very excellence of some sections perforce bestows an unwarranted authority on the rest. Dr. Metzner is a strong advocate of Wolpe's techniques of treatment. It does not appear, however, that this advocacy is based on a critical independent assessment of the evidence, but rather that it stems from a direct acceptance of Wolpe's own statements about his results and those of others. This tendency to accept, uncritically, clinical findings which fit conveniently with theory, appears more than once. Thus on page 23 the belief of several therapists that insight per se is not therapeutic appears a paragraph later as a fully proven observation. Even if Wolpe's therapeutic results are better than conventional methods this is more likely a function of Wolpe than his theories. Certainly the null hypothesis that insight is never therapeutically effective can by its nature never be proven. Indeed most therapists would maintain that insight is often therapeutically valuable.

Though some of the faults of this monograph may be thought important, they are clearly over-shadowed by its virtues—the lost virtues of English learning theorists—brevity, lucidity, and accuracy.

As a vade-mecum to the relationship between treatment and learning theory this is almost certainly the best buy.

ALICK ELITHORN.

CORRESPONDENCE

John Brebner and Ian Gordon write:—

We wish to express our regret to Dr. Mowbray[1] for any misrepresentation of his position which arose from our interpretation of the phrase "signal rate" as in "Furthermore, rate of signal appearance is a common variable in vigilance research that also appears in the reaction time study." to refer to time and not to the signal to non-signal ratio which Mowbray has now pointed out was intended. In this context, "ratio" seems a less ambiguous term than "rate" and we propose to use "ratio" in future.

With respect to Mowbray's comment on the possible role of the distribution of intersignal intervals in our experiment, we agree that these distributions were probably "J" shaped and nearly rectangular with two and eight alternatives respectively. But, with equiprobable items and a constant rate of presentation, randomness, number of items and the form of the distribution of intersignal intervals are interdependent and any alteration of the latter produces a sequence of items which is no longer random. The distributions of intersignal intervals in our experiment derive from random selection and to have structured the sequence of items so as to hold the form of the distribution of intersignal intervals constant between our three conditions would have imposed some constraint upon this randomness.

For our part, we consider it likely that several variables may affect selective response times, these include:—

(i) Signal probability.
(ii) Intersignal interval.
(iii) Number of non-signal categories.
(iv) Discriminability of signals and non-signals.

To these, in the light of Mowbray's unpublished work should, perhaps, be added variations of distribution away from those arising naturally from randomness. As yet, however, experimentation has served to elicit these relevant variables rather than to study their independent effects, and their relative importance is not entirely clear from published data. But, in support of Mowbray's contention that the major determinant is the ratio of signals to non-signals, we have results from three experiments involving manual and vocal selective responses which indicate that where signal probability differs but mean intersignal interval and the number of non-signal categories are held constant, significant differences in selective response times obtain. Our results show some slowing of selective response times when the number of non-signal categories is increased although, on the whole, this effect does not reach a significant level.

In summary then, there seems to be no dispute between Mowbray and ourselves with regard to our respective experimental observations. Moreover, the full extent of our agreement with his statement that, "If responses occur relatively often with respect to no responses, then expectancies will be high, latencies low, and vice-versa." may be seen from a previous statement by one of the authors, jointly with Dr. Szafran, in relation to ageing, that increasing the frequency of signals to be responded to. "May serve . . . to improve (the) selective responsiveness of (the) central mechanisms to a high information signal."[2]

REFERENCES

1. MOWBRAY, G. H. (1962). Some remarks concerning Brebner and Gordon's paper "Ensemble size and selective response times with a constant signal rate." This *Journal*, **14**, 117–8.
2. BREBNER, J. and SZAFRAN, J. (1961). The "psychological refractory phase" and ageing. *Nature, Lond.*, **190**, 195–6.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIV 1962 Part 4

REVERSAL OF LEARNING IN OCTOPUS AND THE EFFECT OF REMOVAL OF THE VERTICAL LOBE

BY

J. Z. YOUNG

From the Department of Anatomy, University College, London, and Zoological Station, Naples

Experiments are described on reversal of learned responses in the octopus. It is shown that normal animals are able to reverse learned responses to vertical and horizontal rectangles shown successively but that animals lacking vertical lobes are unable to reverse such a discrimination. These, and related, results are discussed in relation to the hypothesis that learning consists in appropriate conditioning of cells in the optic lobes that respond to particular visual features but that the vertical lobe plays an important part in making the memory effective, perhaps by generalising the conditioning to optic lobe cells other than those directly stimulated in the process of learning.

INTRODUCTION

Study of the capacity of an animal to learn to reverse its response to a given figure or pair of figures provides a valuable means of investigation of the learning mechanism by which representations are set up in its memory. Each representation may be defined as consisting of a set of nervous connections that ensures a particular response to a given situation (see Young, 1960*b*). After a representation has been established ensuring attack at a certain visible figure by an octopus we can give the animal shocks if attacks are made and thus ultimately establish a representation ensuring the opposite response. Does this new representation involve a distinct set of connections, or is it a modification of the old? Should we think of two sets in competition with each other?

After removal of the vertical lobe the setting up and use of representations in the memory is impaired (Boycott and Young, 1957). In particular it has been clear since the earliest experiments on the subject that animals without vertical lobes cannot readily set up and maintain representations that reverse any that have been previously established (Boycott and Young, 1958). Extinction as a result of non-reinforcement still occurs (Young, 1959) but the animals cannot learn to respond negatively to a previously "positive" stimulus or vice versa.

The present experiments with operated animals were designed to enable comparison between the performance by animals learning discriminations for the first time with similar groups learning after reversal.

METHOD

Octopuses were kept in opaque tanks with good water circulation as already described (Young, 1958). They were trained by showing opaque plastic figures moving up and down at about 3 times a sec. on a transparent plastic rod. Food was given as small (1 to 2 gm.) pieces of fish, placed near to the figure on a wire. Shocks were 11 v. A/C administered by a pair of electrodes introduced when necessary and touched against the body or arms. Each figure was shown for 15 sec. If after that time no attack had been made at the "positive" figure the octopus was fed.

The figures used were horizontal and vertical rectangles (10 × 2 cm.) and black and white circles (5 cm. diameter). The members of the pairs have not the same stimulating value for an octopus (Young, 1958; Muntz and Sutherland, 1959). When moved up and down the vertical rectangle is the more likely to be attacked. The white circle is more likely to be attacked than the black. Reversal of learned responses is therefore more difficult if the animals are first trained to attack the vertical or the white figure than vice versa.

The animals were trained in groups, usually of 5 or 10, the experimenter passing round from tank to tank. The interval between trials with each animal was five minutes.

The two figures were shown in strict alternation and shocks were given for attacks at one, food for attacks at the other. Each day there were two sessions of ten trials each (five with each figure). The sessions were given in the morning and evening, with about six hours between. Operations and assessment of amount of vertical lobe removed were as indicated by Boycott and Young (1957).

Wherever possible groups of animals that were to be compared were trained on the same days, so that their conditions were similar. This was not possible for all the comparisons, but conditions did not vary greatly. Temperatures were between 21 and 25° C and the water circulation was abundant.

The following groups of animals were used:—

Normal animals

Group 1

Five animals. Trained first to attack a vertical rectangle but not a horizontal one (7 days, 70 paired trials) and then after an interval of 3 days the reverse (80 paired trails).

Group 2

Ten animals. Trained to attack a horizontal rectangle but not a vertical one for 80 paired trails. This serves as a control for Group 1.

Animals without vertical lobes

Group 3

Five animals trained exactly as Group 1 and at the same time as that group.

Group 4

Eight animals trained to attack a horizontal but not a vertical rectangle for 9 days (90 paired trials). From the 10th day the horizontal rectangle alone was then shown and shocks given when it was attacked (120 trials).

Group 5

Eleven animals trained to attack a black circle but not a white one for 6 days (60 paired trials) and then from the 7th day the reverse (60 paired trials).

RESULTS

Group 1. *Reversal by normal animals. Vertical rectangle originally positive*

Five normal animals proved able to make the relatively difficult change from vertical to horizontal as the positive figure (Table I, Figure 1). The learning in the new direction was slower than the original learning, but this would be expected from the fact that there is a greater tendency to attack the vertical figure.

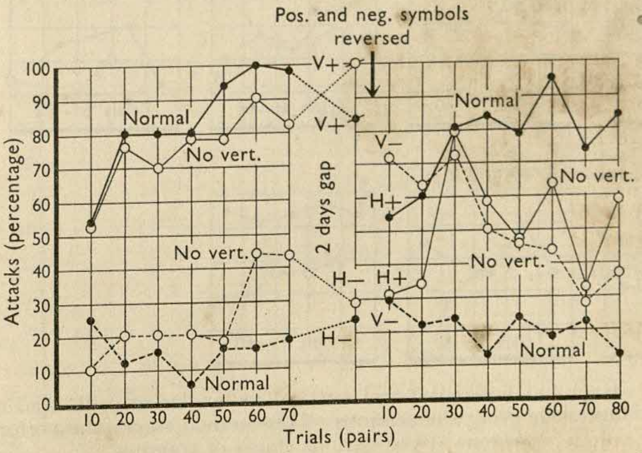
A direct comparison can be made with Group 2, 10 animals trained with the horizontal as positive figure without previous training in the opposite direction. Throughout the training the control series were usually more accurate. The difference was not significant, however, there were 22 per cent. errors in 60 paired trials by the reversed animals, compared with 18 per cent. by those learning for the first time. The level of errors reached after 60 trials by the animals learning the reversed discrimination was about 11 per cent. as against 5 per cent. without reversal and the accuracy showed some wide fluctuations by the individual reversed animals.

TABLE I

Group	*Conditions*	*No.*	*Attacks*		*Errors per cent.*	$\frac{x}{x+y} \times 100$	*Time to Attack+*	
			+ x	− y			*Mean sec.*	*S.D. sec.*
1	Normal V+	5	244	46	17	84		
2	Normal H+	10	516	135	18	79	5·8	3·0
1 (Reversal)	Normal H+ after V+	5	228	62	22	79	7·8	4·2
3	No vert. V+	5	225	68	24	77		
4	No vert. H+	8	313	267	45	54	7·8	3·9
3 (Reversal)	No vert. H+ after V+	5	157	173	53	48	11·1	3·9
5	No vert. B+, W−	11	292	160	40	65		
5 (Reversal)	No vert. W+ after B+	11	125	176	54	41		

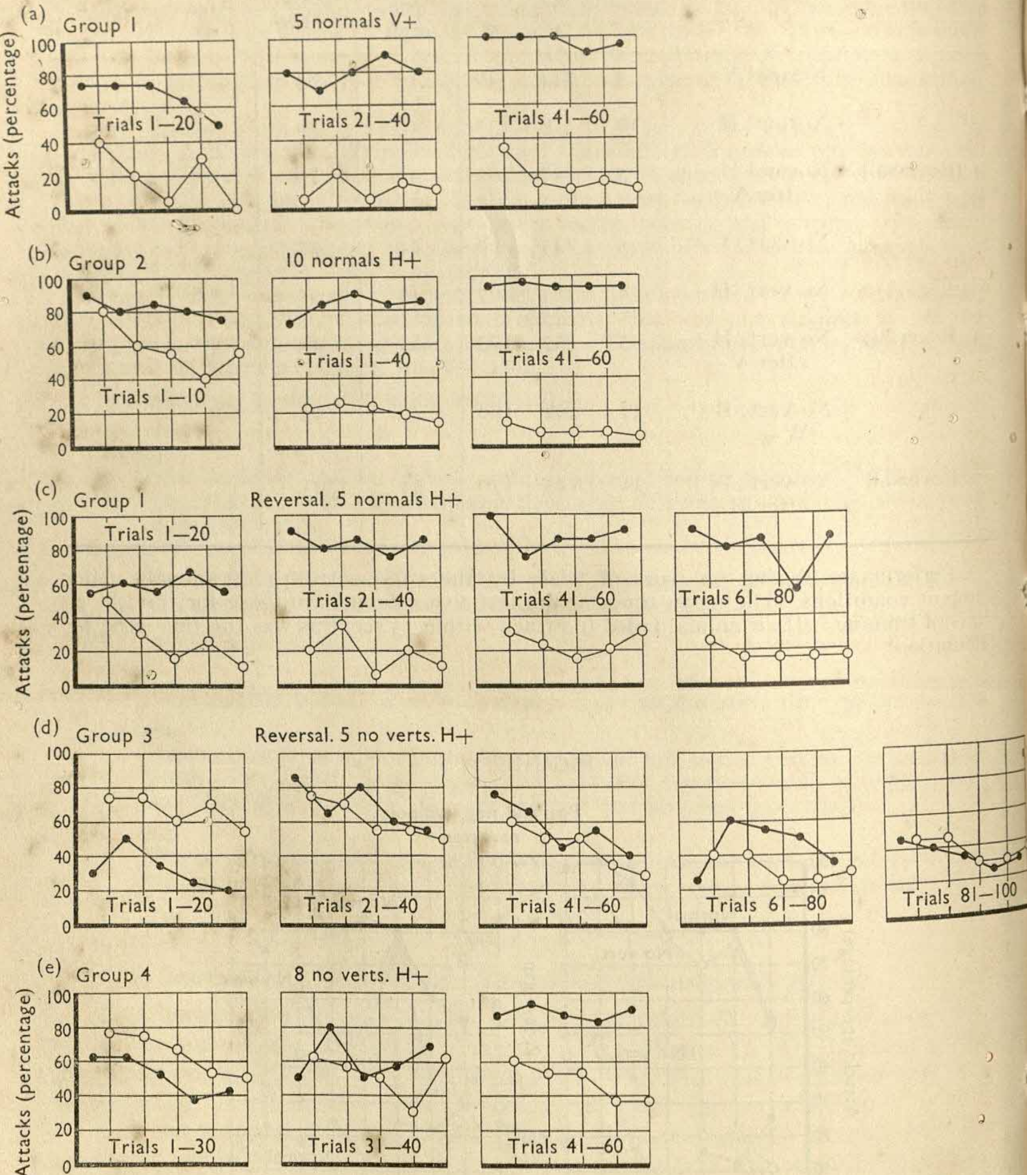
Performance during 60 pairs of trials by the various groups of animals under different conditions. The mean times in the last two columns are those for the last 20 trials of training. (If an animal failed to attack within 15 sec. this was the time used for calculation).

FIGURE 1



Percentage of attacks at positive and negative figures by five normal and five no vertical animals (94 per cent. removed) before and after reversal. Filled circles normal animals, open circles operated animals, Groups 1 and 3. In all figures symbols crossed by a line (*e.g.* ∅) indicate trials without food or shock.

FIGURE 2



Percentage of attacks at horizontal and vertical rectangles at each trial of the session, averaged over successive groups of sessions. This method shows the performance within sessions at the various stages of training.

The times taken between showing the positive figure and completion of the attack also oscillated considerably for each animal and were often longer, not only than those during the initial learning by the same animals, but also than those for the control series. The mean times over trials 41–60 were 7·8 sec. for Group 1 (with reversal) and 5·8 for Group 2, but the difference is not significant. The mean standard deviations of the attack times, however, were 4·2 for the reversal group and 3·0 for Group 2, a difference that is significant between 0·05 and 0·01 level.

Even after 80 reversed training trials the attack at the new positive figure was often markedly cautious. When the negative figure (vertical rectangle) was shown the octopuses usually approached it. They sometimes stretched out an arm towards it but often waved the arm in some other direction, such as towards the surface, or the jet of water running into the tank. (Would ethologists call this a displacement activity?) The representation by which the vertical rectangle elicited a positive response, therefore, still persisted, but was in conflict with a new one preventing attack. The outcome was a movement in some other direction.

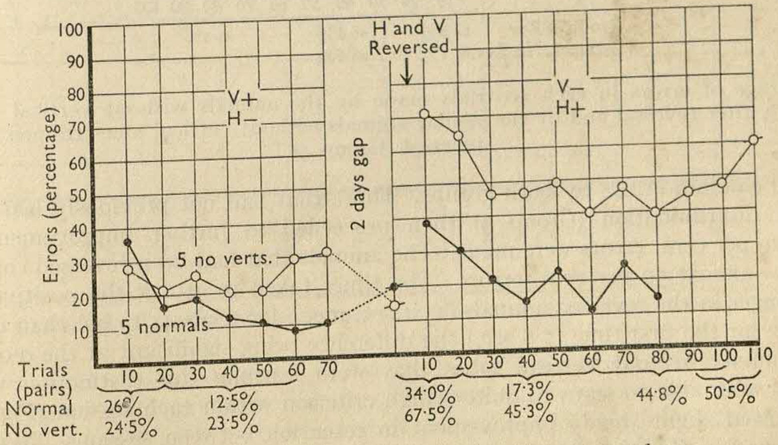
When the pattern of responses for each of the ten trials of the nine sessions is analysed, animals learning for the first time show learning within each session, even in the later sessions of the series (Figure 2b). Animals learning the reversed discrimination show a less consistent pattern, performing little better at the ends of the sessions than at their beginning (Figure 2c). All the five animals learning the reversal showed the same general pattern.

Reversal after vertical lobe removal

Group 3. Vertical rectangle as original positive

A set of five octopuses from which the vertical lobes had been removed three days previously at first showed clear signs of learning when trained to attack the vertical but not the horizontal rectangle (Figure 1), but gave no further improvement after the first ten trials. They were trained on the same days as the normal animals of Group 1 learning the same discrimination. The operated animals were less accurate in discrimination except slightly over the first ten trials. Even after 70 trials they were making 30 per cent. errors, compared with less than 10 per cent. by the normal animals. The times taken to attack were also greater in the operated

FIGURE 3



Percentage of errors in each 10 trials. Groups 1 and 3.

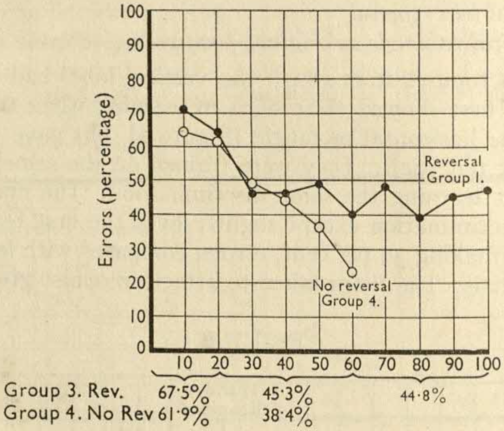
animals (Table I). The apparently low variance of these times in the operated animals is an artefact due to their slowness. Since on many occasions no attack was made within 15 sec. the values cluster around this figure. A study in which they were left for longer than 15 sec. would probably show greater variance in the operated animals.

Thus representations ensuring correct response could be set up by these operated animals. Moreover these were well retained when the animals were left for two days without training. Tested without reward they then performed approximately at the same level as the normal animals (15 to 20 per cent. errors, Figure 3).

However, they proved quite unable to reverse this response. When the rewards were changed they at first continued to attack the vertical but not the horizontal figure. Then after twenty trials they began to attack both figures frequently. Gradually the attacks became less and less numerous until only about 3/10 attacks were made at either rectangle. The proportion of errors became a little less than 50 per cent. (Figure 3).

We can compare these animals with the control series of eight octopuses without vertical lobes trained to attack a horizontal but not a vertical rectangle without previous training in the reverse direction (Group 4). Both groups at the start of training on this discrimination showed "perverse" behaviour, because of the greater tendency to attack the vertical figure (Figure 4). This was soon corrected to the level of random errors by both groups. Those that had not previously learned the reverse discrimination (Group 4) then proceeded to further improvement until about 20 per cent. errors were made. The animals that had been reversed continued to make almost 50 per cent. errors. The times taken to attack the positive figure were longer in the reversed animals (11·1 sec. mean for trials 41 to 60) than in those learning for the first time (7·8 sec.) the difference being significant at the 0·02 level. The animals without vertical lobes that were learning this distinction without reversal were able to learn to quite a high criterion within each session (Figure 2e) and showed slight steady improvement in retention between sessions. Neither of these was possible for the operated animals learning after reversal (Figure 2d).

FIGURE 4



Percentage of errors in each 10 trials made by the animals without vertical lobes of Group 3 after reversal and by the similar animals without vertical lobes trained without reversal (Group 4).

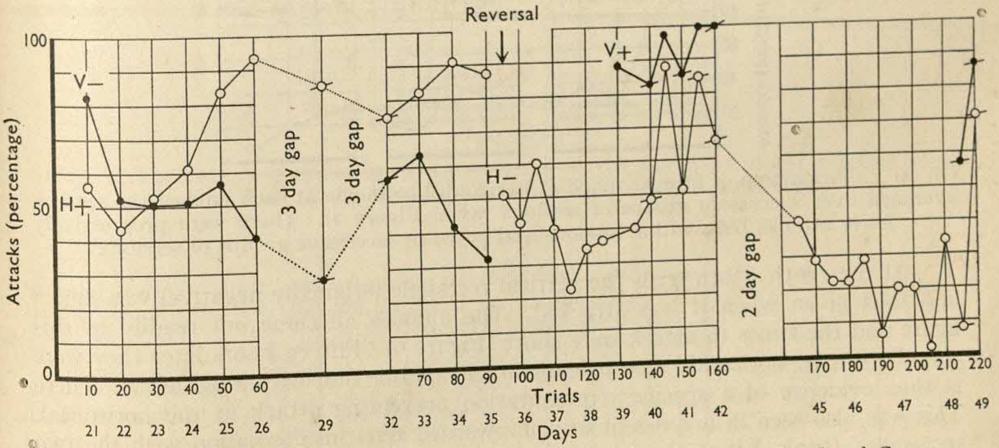
Even after 100 trials they were no better at the ends of the sessions than the beginnings. Any explanation of the functions of the optic and vertical lobes must therefore assume for the present a complete inability of the optic lobes alone to reverse a visual discrimination. It will be interesting to confirm this for other figures.

There were considerable individual differences among the operated animals during reversed learning. One animal did not show the decline in attacks that was seen in the others towards the end of the experiment. This decline must be considered to be a slow learning to associate both figures with shock, without the differentiation that all had proved able to make to some extent during the initial learning before reversal.

Group 4. Reversal with horizontal rectangle as original positive figure.

An alternative method of undertaking a reversal is to train animals to attack one figure and not another and then to reverse the rewards for one only of these, omitting trials with the other. This should be an easier task than reversal to both, nevertheless it proved difficult for animals without vertical lobes to learn in this way not to attack the previously positive figure. The eight animals used had been operated 20 days previously. All had some parts of the vertical lobes remaining, the mean amount removed being 78 per cent. They were first taught to attack a horizontal and not a vertical rectangle, and a moderate level of accuracy was achieved (Figure 5). Tests given after intervals of three and six days showed good retention. After twenty further paired training trials the attempt was then made to reverse the reaction to the horizontal rectangle by showing it for numerous trials and giving shocks when it was attacked. No trials were given with the vertical rectangle at first, the animals being fed at the end of each day in their homes.

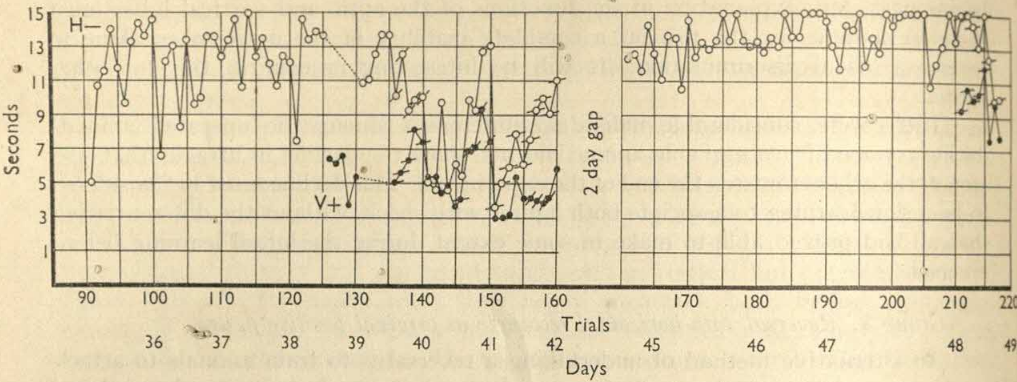
FIGURE 5



Percentage of attacks at the positive and negative figures by the animals of Group 4 (no vertical lobes) before and after reversal. No trials were given during the gaps.

The effect of the shocks was to reduce the tendency to attack to about one in three occasions after 40 trials (Figure 5). However, since the trials were given at 5 min. intervals, in groups of fives, the overall picture gives a false impression. At the beginning of each session many animals would attack and the time of attack

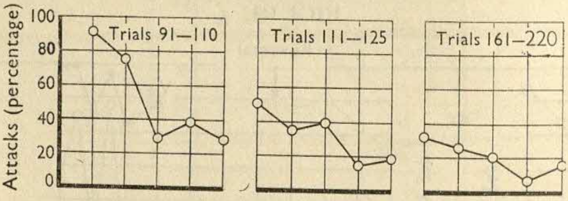
FIGURE 6



Mean times to attack at each trial by the animals of Group 4 while learning not to attack the horizontal rectangle, a previously positive figure, and, later, to attack the vertical rectangle, (previously negative). Open circles horizontal, filled circles vertical rectangle.

was short (Figure 6). Then as the session progressed attacks would become slower and one after the other of the animals would fail to attack at all within the 15 sec. for which the figure was shown. This fall in attacks throughout each session is brought out by grouping several successive sessions together (Figure 7). Besides the learning not to attack within each session there was obviously a slow development of representations that prevented attacking even at the beginning of each session. The mean percentage of attacks over trials 101–135 was 47 per cent. and for trials 180–210 21 per cent.

FIGURE 7



Group 4. Proportion of attacks made at horizontal rectangle at each trial of the session, averaged over successive groups of sessions (as in Figure 2). There were progressively fewer attacks both within sessions and between successive groups of sessions.

At the 126th–130th trials the vertical rectangle (originally negative) was shown and food given when it was attacked. The animals all came out readily to this figure and the times to attack were short (Figure 6). Fifteen hours later they were tested with the horizontal rectangle again. Few came out, with long delays. There is thus evidence of a specific representation preventing attack at the horizontal. This was also seen in a series of five unrewarded tests in alternation with the two rectangles (trials 136–140). Then at the 141st–145th trials the rectangles were shown alternately, shocks being given for attacks at the horizontal and food for those at the vertical. Attacks were made on almost every occasion at which either figure was shown (Figure 5). The representation preventing attack was not strong enough to overcome the stimulus to attack that is provided by the food. Next, no rewards were given for attacks at the vertical and the attacks at the horizontal decreased, but began again when food was again given with the vertical (151–155).

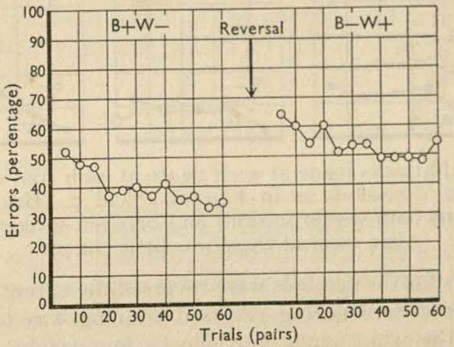
After an interval of 2 days many further groups of trials with the horizontal rectangle then followed, at which shocks were given for attacks. There was a further slow but not very steady decrease in attacks (Figure 5). Each animal was liable to come out from time to time to attack. There was still evidence of learning within each session (Figure 7), but there were also fewer and slower attacks than previously at the first trial of each session (Figure 6).

Tests without reward with the vertical and the horizontal rectangle then again showed fewer and slower attacks at the horizontal (trials 211–215), but when food was given as reward for attacks at the vertical, the attacks at the horizontal, shown at intervening trials, again became more frequent (trials 216–220, Figure 5 and 6). In spite of all the shocks no representation strong enough to prevent attacks at the previously positive horizontal figure under these conditions had been set up.

Group 5. *Reversal of response to black and white circles after vertical lobe removal*

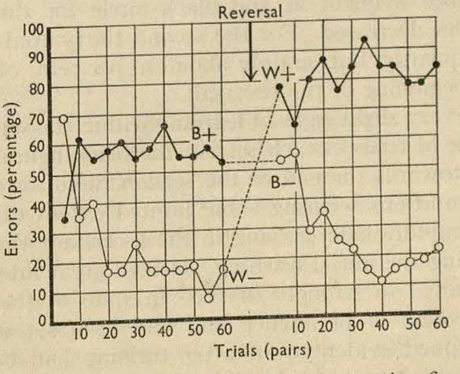
The eleven octopuses of this group were trained 31 days after removal of the vertical lobes (mean of 88 per cent. removal) to attack a black circle but not a white one. The figures were shown successively. The direction of training was then reversed (Figure 8). During the initial training the attacks at the white circle declined to about 15 per cent. while those at the black circle were made at only

FIGURE 8



Group 5. Percentage of errors in each five trials by the animals without vertical lobes learning first to attack black but not white and then the reverse.

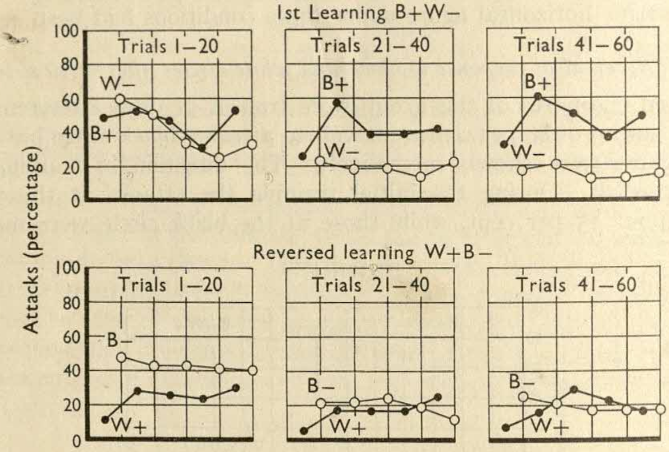
FIGURE 9



Group 5. Percentage of attacks with the positive and negative figures shown separately.

slightly over half the occasions (Table I and Figure 9). This failure to attack was not due to ill-health or injury to underlying parts of the brain. Before the training began the animals were all persistent attackers of crabs. The poor performance with this easy discrimination between black and white is another example of the difficulty of setting up representations in the absence of the vertical lobes. The phenomenon of learning within each session was present (Figure 10). Throughout each of earlier sessions attacks at both figures declined, those at the negative figure slightly more than at the positive. In later sessions behaviour was more consistent throughout, with fewer attacks at the negative.

FIGURE 10



Group 5. Proportion of attacks made at each figure at each trial of the session averaged over successive groups of sessions as in Figures 2 and 7. Before reversal there was progressive improvement both within sessions and between successive groups of sessions. After reversal there was little change.

When the rewards for the symbols were reversed the extent of the initial learning became clear. The procedure during reversed training was to give food on every occasion of showing the white circle and shocks for attacks at the black one. If the octopus did not attack the white circle within 15 sec., a piece of fish was introduced on a wire, and the circle and the fish together were pushed towards the home. All the animals readily came forward to receive the food. Nevertheless attacks remained more frequent at the black circle for thirty trials (Figure 9). Attacks at both circles decreased. For the second thirty trials the two figures were attacked equally frequently, but on only about 20 per cent. of trials, against nearly 50 per cent. at the beginning of the reversal.

There were only very slight signs of learning within the separate sessions (Figure 10). At the first pair of trials of each session the black figure was always the more often attacked, but towards the end of the session there were more attacks at the white. Thus representations securing some limited extent of correct response can be set up, but they rapidly fade, leaving in the ascendant the representations that had been set up during the initial learning. It is of great interest that these initial representations remained so strongly marked in spite of the fact that they were contrary to the direction of preference and had been set up after vertical lobe removal. They remained evident even after training had been against them for 60 paired trials, extending over six days.

There were marked individual differences. The worst performers were two animals that had been among the better learners of the original discrimination. As might be expected their especially large number of errors was due to their persistence in attacking the black circle, which is reasonable if they had acquired stronger representations to do this initially. Indeed all those animals that had shown some signs of initial learning made more than 50 per cent. errors during reversal. Those giving approximately random responses before continued to do so afterwards and thus now appeared as the "better" performers.

The only representation set up during the reversed training was thus one tending to prevent attack at either figure. At the end of the experiment it was necessary to show that this failure to attack was not due to some deficiency that had developed during the long period of feeding with fishes. Five trials were therefore given at which the reward with the white circle was a live crab, the crab being given near the circle at the end of 15 sec. if there had been no attack. During this set of five trials with the white and five with the black circle there was only one attack while the crab and the circle were viewed together at a distance (1/55). Yet the octopuses all came out to attack when the crab alone was moved near to them.

On the next day the animals were again tested with the crabs and circles together and not a single animal attacked. Nevertheless when shown crabs *alone* all except one attacked, some in as little as 3 sec. This readiness to attack confirms that the progressive failure to attack the circles was dependent on a representation preventing attacks that had been set up in the memory as a result of the shocks. This result thus agrees closely with that shown by the attempt to reverse the responses to rectangles (p. 198).

DISCUSSION

Reversal by normal animals

In normal octopuses it is readily possible to set up representations in the memory ensuring responses opposite to those originally formed. If the original representations were well established, however, clear signs of them could still be seen even after 80 pairs of reversed training trials. This suggests that the representations associating a given figure with food or shock are not readily destroyed. The reversed response is probably produced by new sets of neural connections rather than by attaching the opposite rewards to the old representations of the figures. These results are thus in agreement with the hypothesis that during learning in an octopus the responses produced by particular detector cells (*e.g.* for vertical and horizontal stimulation) become limited and are not thereafter readily changed (Young, 1960*a*).

Reversal after vertical lobe removal

The difficulties involved in setting up reversed representations are greatly exaggerated after vertical lobe removal. When rewards are reversed improvement from the initial behaviour occurs up to the 50 per cent. level of errors, but not beyond this. The animals tend even more than normals to cease attacking either figure. This behaviour appeared clearly in two distinct sets of animals with different discriminations (horizontal and vertical rectangles and black and white circles).

The processes involved in reversal by animals without vertical lobes appear in the results of reversing responses to a "positive" figure only. When shocks are given for attacks at a figure that has previously been rewarded with food the animals gradually come to attack it less often. If they are then stimulated to attack by feeding, however, they again begin to come out to attack. The representations associating the figure with food were thus still present and assume dominance when the general tendency to attack was raised.

There is thus clear evidence from these experiments that it is very difficult to set up reversed representations of visual situations after vertical lobe removal. Yet it is not easy to specify clearly the nature of the activities by which these lobes assist in memory formation. The failure to establish adequate representations suggests that they participate in the process of setting up these representations in the memory. This they might do either by providing extra units that can be conditioned to produce appropriate responses, or by re-excitation of the optic lobes to provide access to further cells there that are not influenced by the direct presentation from the retina, and thus to generalise the conditioning throughout these lobes (Young, 1960*a* and *b*).

On the other hand the fact that approximately correct reversed responses to a new negative figure may be obtained provided no food is present (p. 200) suggests that the vertical lobes are involved in ensuring the proper effectiveness of representations in the control of behaviour, by some type of damping or restraining of the effects of "unconditional" stimuli. It may be a condition of all learning systems that immediate responses should be inhibited in this way, allowing the learned representations to have their effect. The failures of animals without vertical lobes are then due to a lack of inhibition of unlearned responses.

This conception is attractive and may perhaps be extended to cover all the known effects of vertical lobe removal, for instance the variability of behaviour and of time of response (see also Maldonado, 1963). Perhaps most difficult of all to understand is the progressive failure to attack by operated animals when "attempting" to learn a reversed discrimination.

It is clear that representations can be set up outside the vertical lobes and yet equally clear that these lobes play a fundamental part in the operation of the memory system as an effective agent for the control of behaviour.

SUMMARY

Normal octopuses are able to reverse learned responses to vertical and horizontal rectangles shown successively. Learning in the new directions is somewhat slower than the original learning. Even after long reversed training responses to the new positive figure were somewhat slow and cautious. The old representations thus remain in the memory but are partly obscured by new ones working in the opposite directions.

After removal of the vertical lobes discrimination of vertical and horizontal rectangles was learned to a lower standard of accuracy than in normal animals. Responses remained slow and variable even after long training.

Five octopuses without vertical lobes proved unable to reverse a discrimination between vertical and horizontal rectangles. They gradually ceased to attack either figure.

Eleven octopuses without vertical lobes showed similar incapacity to reverse their responses to black and white circles.

Eight octopuses without vertical lobes after training to attack a horizontal but not a vertical rectangle were given repeated shocks for attacks at the horizontal. They gradually learned not to attack, at first learning only within each session of trials, but later with greater reliability. As soon as food was given, however, they again came out to attack the horizontal rectangle. The representation tending to produce such attacks was thus still present.

These results are discussed in relation to the hypothesis that learning consists of appropriate conditioning of cells in the optic lobes that respond to particular visual features (black or white, vertical or horizontal). The vertical lobe certainly plays an important part in making the memory effective. Some results suggest that it does this by generalising the conditioning to optic lobe cells other than those stimulated at the

moment of training. Other results suggest that the lobes have a damping influence, preventing the swamping of the effect of representations in the memory by innate tendencies such as are produced by feeding.

The work reported in this and the following paper was done at the Zoological Station at Naples and it is a pleasure to thank the Director, Dr. P. Dohrn and his staff for their help. My thanks are also due to Dr. J. de C. Downer, J. and N. Mackintosh and Dr. N. S. Sutherland for discussions of the work and especially to Miss J. Parriss and Miss M. Beaumont for help in the preparation of the paper.

REFERENCES

BOYCOTT, B. B., and YOUNG, J. Z. (1957). Effects of interference with the vertical lobe on visual discrimination in *Octopus vulgaris* Lamarck. *Proc. roy. Soc., B.*, **146**, 439–59.

BOYCOTT, B. B., and YOUNG, J. Z. (1958). Reversal of learned responses in *Octopus vulgaris* Lamarck. *J. anim. Behav.*, **6**, 45–52.

MALDONADO, H. (1963). (To be published).

MUNTZ, W. R. A., and SUTHERLAND, N. S. (1959). Simultaneous discrimination training and preferred directions of motion in visual discrimination of shape in *Octopus vulgaris* Lamarck. *Pubbl. Staz. Zool., Napoli*, **31**, 109–26.

YOUNG, J. Z. (1958). Responses of untrained octopuses to various figures and the effect of removal of the vertical lobe. *Proc. roy. Soc., B.*, **149**, 463–83.

YOUNG, J. Z. (1959). Extinction of unrewarded responses in *Octopus*. *Pubbl. Staz. Zool., Napoli*, **31**, 225–47.

YOUNG, J. Z. (1960*a*). The visual system of *Octopus*. (1) Regularities in the retina and optic lobes of *Octopus* in relation to form discrimination. *Nature, Lond.*, **186**, 836–9.

YOUNG, J. Z. (1960*b*). Unit processes in the formation of representations in the memory of *Octopus*. *Proc. roy. Soc., B.*, **153**, 1–17.

REPEATED REVERSAL OF TRAINING IN OCTOPUS

BY

J. Z. YOUNG

From the Department of Anatomy, University College, London

Octopuses were trained by successive presentations to discriminate between two figures by means of a combination of food rewards and electric shocks. This was followed by a series of up to nine daily sets of trials in which reward and punishment were systematically reversed on alternate days. It was found that performance became progressively less accurate with successive reversals until it reached a random level; this appeared largely due to a continual fall in the total number of attacks at the figures. At the same time, such attacks as were made were mainly at the positive figures and there was evidence that performance as judged by this criterion became more accurate during each day and perhaps progressively more accurate on successive days. Octopuses lacking the vertical lobe showed a fall in the total number of attacks but did not show this improvement within each day or over successive days. A theoretical interpretation of the results is put forward.

INTRODUCTION

When octopuses have been made to reverse a learned visual discrimination they long continue to show signs of persistence of the original responses (Young, 1962). This suggests that reversal is not so much a matter of attaching reversed rewards to existing representations of the figures as of building completely new assemblies in the nervous system to represent the new situation. What will then be the effect of subsequent further reversals? Will it be possible to build a number of such representations and will they produce cumulatively more accurate effects or conversely interfere with each other?

When faced with repeated reversal of rewards many mammals show a progressive improvement in performance (Harlow, 1949; Cronholm, Warren and Hara, 1960; and others). This improvement has not in general been found in other animal groups (see Discussion).

In the present investigation the direction of training was repeatedly reversed for discrimination between pairs of figures presented successively. Food and shock were reversed each day. The effects of repeated reversal after training of octopuses to a criterion have been studied by Mackintosh (1962) (see p. 220).

The vertical lobes were removed from one group of animals before the series of reversed training was begun. The difference between the performance of these animals and normals has provided some insight into the mechanism involved in reversal learning.

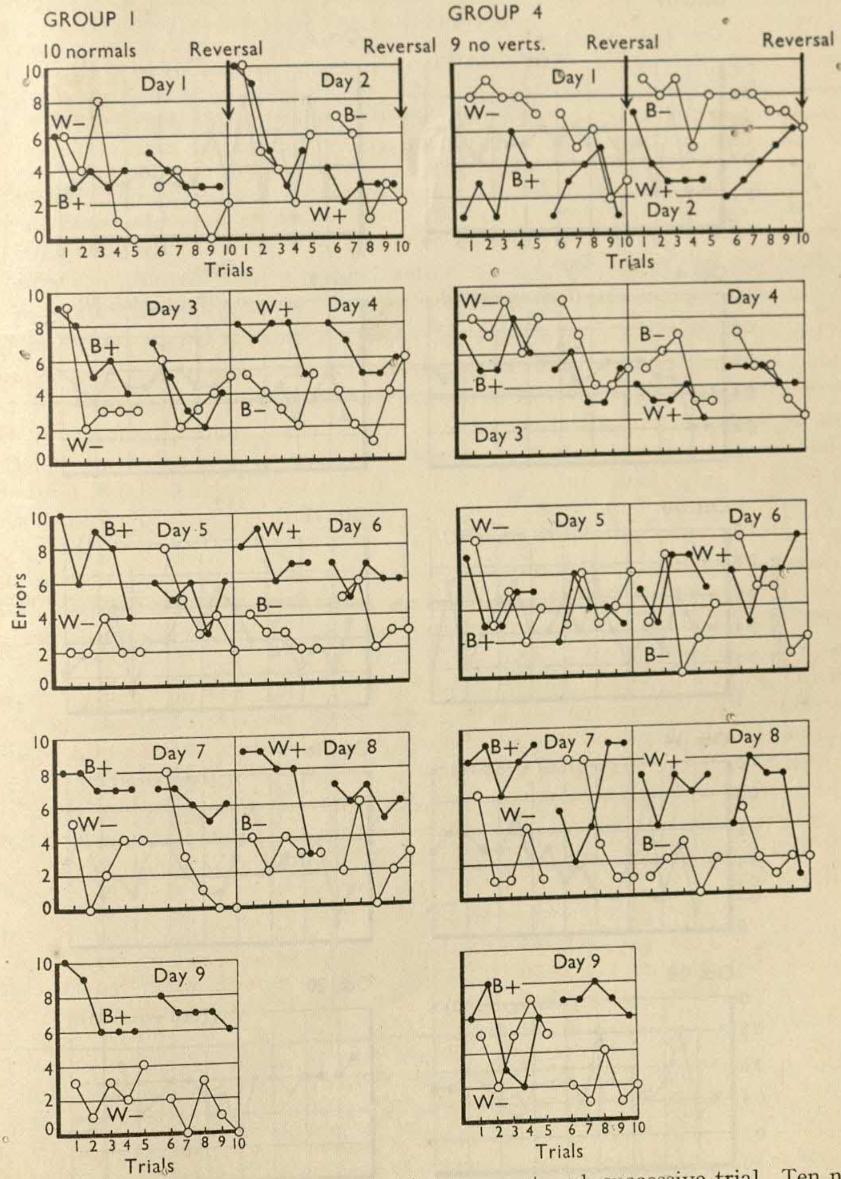
METHOD

The octopuses (*O. vulgaris*) were kept singly in opaque-sided asbestos tanks 100 × 40 × 30 cm. with plentiful sea-water circulation. Normal animals were studied several days after they had been brought from the sea. During these preliminary days they were fed with sardines at the end of the tank distant from the home. The animals ranged in weight from 250–500 gm. Training was with opaque plastic figures, moved up and down at the end of the tank distant from the home (Young 1962). The food reward was pieces of fish (1–2 gm.) placed near to the figure on a wire. Food was given at every trial with the positive figures (unless stated otherwise); if the octopus had not attacked after the figure had been presented for 15 sec. the food was introduced and moved with the figure towards the octopus until it was taken. Shocks (11 v. in some experiments, 6 v. in others) were given only if the animal attacked the negative figure.

Trials with the figures were given in regular alternation of positive and negative at five-minute intervals, there being two sessions each of ten trials (five positive and five negative) each day. The positive figure for the day was shown at the first trial of each session. One session was given each morning and the other in the evening, with about six hours interval between.

At each trial the time taken between insertion of the figure and seizure of it was noted with a stop-watch. A record was kept of the occasions at which the animal left its home and made an "approach" to the figure without attacking it.

FIGURE 1



Groups 1 and 4. Number of animals making errors at each successive trial. Ten normal and nine no vertical animals during initial learning (day 1), and eight reversals (days 2–9) with black and white circles.

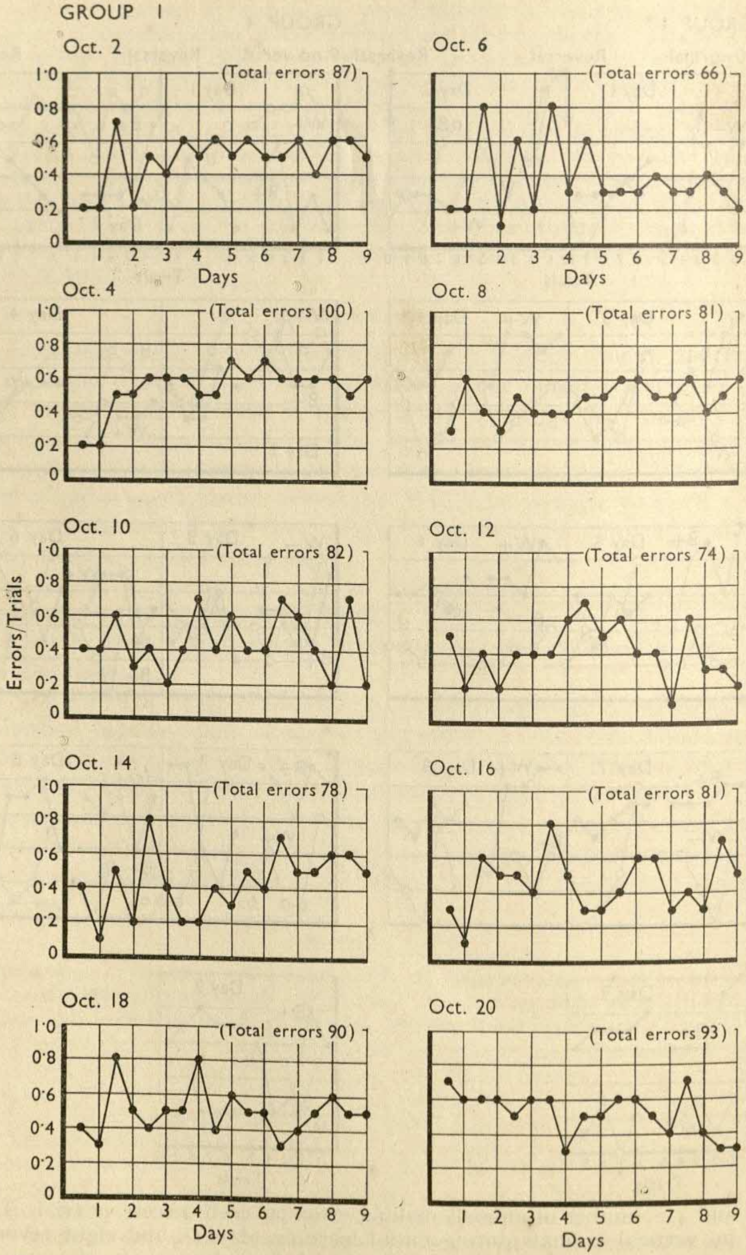
Operations on the vertical lobe and assessment of the amount of tissue removed were as described by Boycott and Young (1957).

Four groups of animals were used:—

Group I

Ten normals trained with black and white circles, diameter 5 cm. Food and shock rewards reversed each day for nine days.

FIGURE 2



Group I. Proportion of errors to trials made by individual normal octopuses during initial learning, and eight reversals with black and white circles.

Group 2

Eight normals trained with horizontal and vertical rectangles, 10 × 2 cm. Food and shock rewards reversed each day for six days.

Group 3

Eight normals trained with squares 5 × 5 cm. shown with base parallel to the floor ("square") or set on edge ("diamond"). Food and shock rewards reversed after each two days, total of four reversals.

Group 4

Nine animals without vertical lobes trained with black and white circles alongside Group 1 and reversed in the same way as that group.

The figures used in these pairs are not equally effective in stimulating an octopus to attack. White is more likely to release an attack than black and vertical than horizontal (when both are moved vertically). The "diamond" may be more stimulating than the "square" (Sutherland, 1960), but the difference is small.

RESULTS

Normal Animals

Group 1. *Black and white circles*

Initial training

These ten animals mostly learned within the first day to attack the black but not the white circle (Figure 1). There were altogether 39 per cent. errors in the first session and 29 per cent. at the second (Table I).

TABLE I

	Attacks		$\frac{x}{x+y}\cdot 100$	$\frac{x}{x+y}\cdot 100$ $B+$	$\frac{x}{x+y}\cdot 100$ $W+$		Errors per cent.	
	x +	y −					$B+$	$W+$
	1	2	3	4	5	6	7	8
R_0	30	19	61	67		39	35	
	32	11	75			29		
R_1	18	27	40		54	59		47
	35	19	65			34		
R_2	18	20	47	54		52	47	
	29	20	60			41		
R_3	14	19	42		48	55		52
	19	17	53			48		
R_4	13	12	52	52		49	49	
	24	22	52			48		
R_5	13	14	48		49	51		51
	19	19	50			50		
R_6	13	15	46	54		52	48	
	19	12	61			43		
R_7	16	16	50		55	50		47
	19	13	59			44		
R_8	13	13	50	60		50	46	
	15	6	71			41		

Group 1. Ten octopuses. Initial learning (R_0) and eight reversals of training with black and white circles. Total of fifty trials with each figure at each session. Columns 1 and 2 show actual numbers of attacks at positive (x) and negative (y). Column 3 shows the proportion of attacks at the positive at each session as a percentage $\left(\frac{x}{x+y}\times 100\right)$. Columns 4 and 5 give the same proportion but for each entire day, column 4 when black was positive and 5 when white positive. Columns 6–8 give the percentage errors, 6 for each session and 7 and 8 for the entire days, with black and white positive separately.

First reversal

The effectiveness of the representations set up during the 20 trials of the first day is shown by the fact that on the next day during the first two trials with the symbols reversed the animals made 100 per cent. errors, the white circle was never attacked, the black was attacked by all ten animals. However there was rapid improvement in response to *both* figures, even during the first few reversed trials (Figure 1). This continued during the second session of the day, during which there were only 34 per cent. errors. There was a greater tendency after reversal to attack the negative, now the black circle. This is especially remarkable since this is the colour that is intrinsically less likely to provoke an attack.

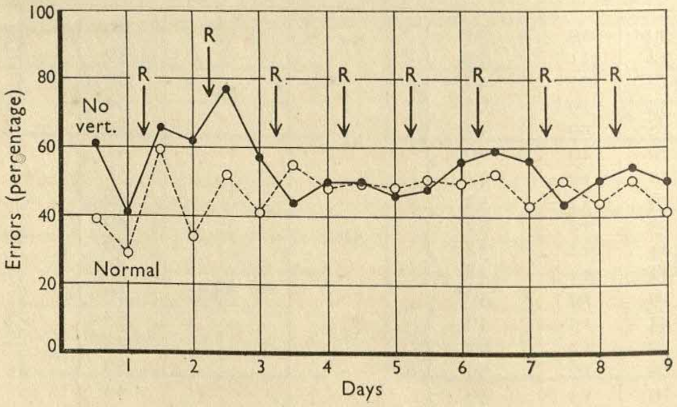
The final level reached at the end of the second day was 5/20 errors at the last trials with the black and white circles, the same as at the end of the first day. Clearly reversal can take place readily after this small amount of initial training.

There was little agreement between the individual performances in the two sets (Figure 2). The correlation coefficient between the errors made in the second ten trials of the two sets was only 0·33.

Second and further reversals

Reverting on day 3 to the original discrimination with the black circle as positive figure the animals at first did not attack it. That is they showed the effect of representations carried over from the previous day (Figure 1). However their performance was less "perverse" than it had been after the first reversal. They learned to some extent to react correctly during the day, but in the second session of trials they made appreciably more errors than on the first two days (Figure 3). During the last two trials of the day there were 9 errors in 20 trials.

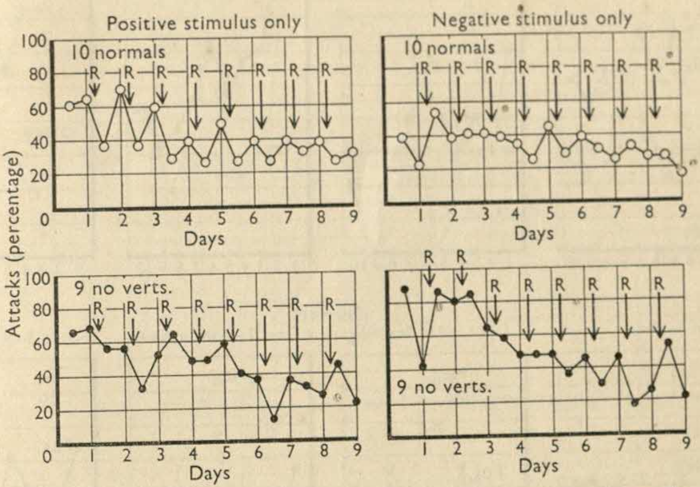
FIGURE 3



Groups 1 and 4. Mean errors at each session of ten trials (five positive, five negative) by ten normal octopuses (open circles) and nine without vertical lobes (filled circles). Initial learning (day 1) and eight reversals (days 2–9). Black and white circles.

With the third reversal there was still less indication during the early trials of the effect of the previous day's training and very little improvement during the day. From here onwards with successive reversals the animals made about 50 per cent. of errors at both the first and second sets of trials of each day (Table I and Figure 3). From the 6th day (5th reversal) food was given to every animal at the first positive trial of the session but thereafter only if there was an attack at the positive figure. There were still slight signs of learning during each day. Thus the positive figure was regularly attacked slightly more often at the second session each day than at the first (Figure 4). The responses to the negative figure were irregular. Attacks became less and less frequent whichever figure was shown.

FIGURE 4



Groups 1 and 4. Percentage of attacks at positive and negative colours for each set of five trials by normal and no vertical octopuses. Initial learning (day 1) and eight reversals (days 2–8).

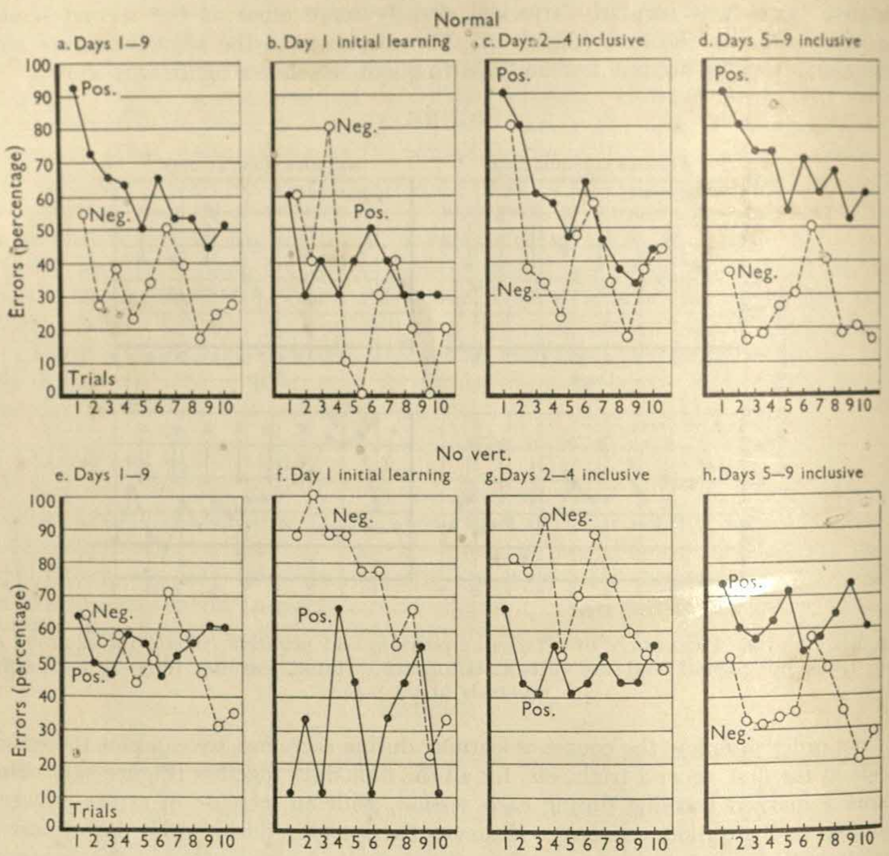
In order to follow the course of learning during each day, we can plot the errors made at the first, second trials, etc. for all the nine days together (Figure 5a). This shows a marked learning during each session, with an increase of errors between morning and afternoon sessions. However the learning during each day occurred mainly in the earlier days of the series, as appears if we plot separately the responses for the initial learning (5b), for the first three reversals (5c) and the rest (Figure 5d).

The clearest change throughout the series was the increasing tendency not to attack (Table I). Attacks became more and more cautious and many approaches were recorded. The times taken to attack the positive figure became in general increasingly longer, although they declined slightly during each day (Figure 6).

In order to allow for the declining tendency to attack, the results have also been calculated to show for each session the proportion of attacks at the positive figure to all attacks (Table I and Figure 7). This suggests that there is actually an *improvement* in performance towards the end of the series. At the end of each day the few attacks that were made were nearly all at the positive figure. This tendency of course was becoming more prominent as attacks became fewer. It suggests that some learning was occurring each day but only became able to show itself against a declining tendency to attack (p. 221).

Considering individual animals, wide differences were seen (Figure 2). The most efficient animal was number 6 with only 66/180 errors. This is a remarkable record, especially considering the large number of "mistakes" made after each of the earlier reversals. This animal showed marked reversal four times and then a decreasing number of attacks at either figure. Other animals showed clear signs of reversal on only the first one or two occasions or not at all (numbers 4 and 20). For several of

FIGURE 5



Groups 1 and 4. Percentage of errors with positive and negative figures at each successive trial throughout the day, added together for days 1–9 in *a* and *e*, day 1 alone in *b* and *f* (initial learning), days 2–4 in *c* and *g*, and days 5–9 in *d* and *h*.

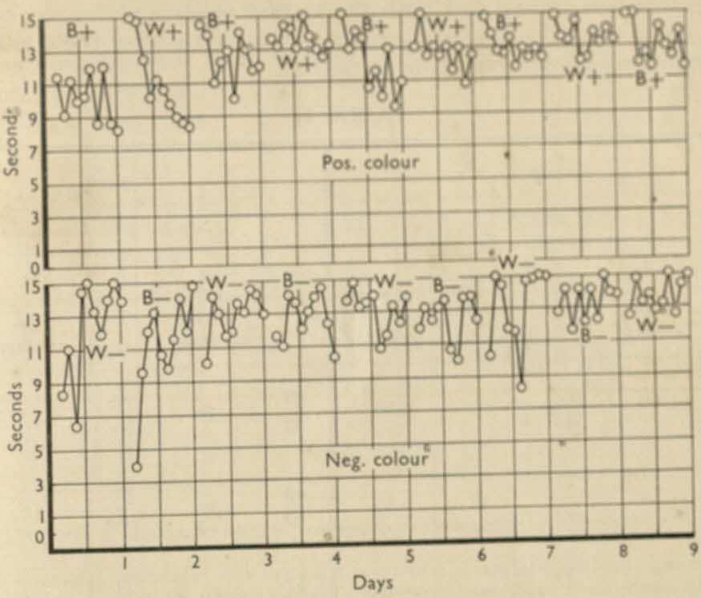
the animals there was a more or less consistent performance, in the sense of numerous or few errors throughout the series, but in others this fluctuated considerably. Correlations between error scores at successive reversals were all positive, except one, and four of these correlation coefficients were significant at the 5 per cent. level, suggesting that there are consistent differences in learning capacity.

Group 2. *Repeated reversal with horizontal and vertical rectangles*

Four octopuses were first trained with the vertical rectangle positive and four with the horizontal positive. Food was given at each positive trial and 6v. shock for each attack at the negative figure. The direction of training was reversed each morning.

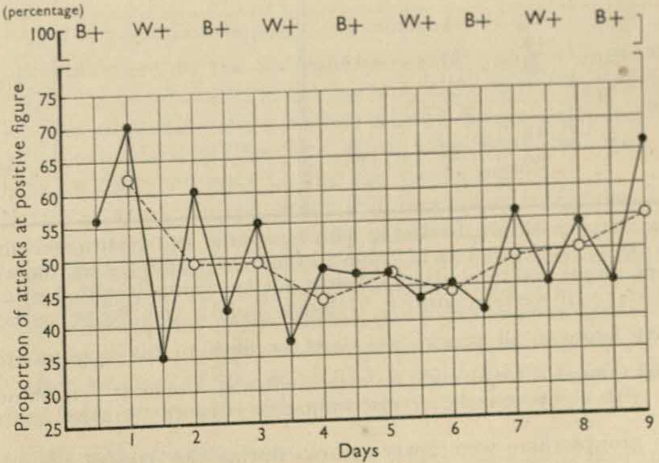
Six reversals were given (Table II). The performances after the odd-numbered reversals were all much worse than after the even-numbered ones, showing the persistence of the effect of the original training. Indeed after the later odd-numbered reversals the discriminations were not learned at all, the scores for these days being at less than the 50 per cent. level.

FIGURE 6



Group 1. Mean times recorded by ten normal animals for attacks at the positive and negative figures. Where no attack was made within 15 sec. this was the time used in the calculation. Points on the 15 sec. line therefore mean that no animal attacked at that trial.

FIGURE 7



Group 1. Proportion of attacks at the positive figure to all attacks ($\frac{x}{x+y} \times 100$ of Table I) plotted for each session during initial learning and eight reversals. The open circles show the mean value for each day.

The effect of the preference for the vertical rectangle was marked (Figure 8). The group that was trained first with the vertical positive performed better than the other group after the even-numbered reversals, worse after the odd-numbered. The tendency was however the same in both.

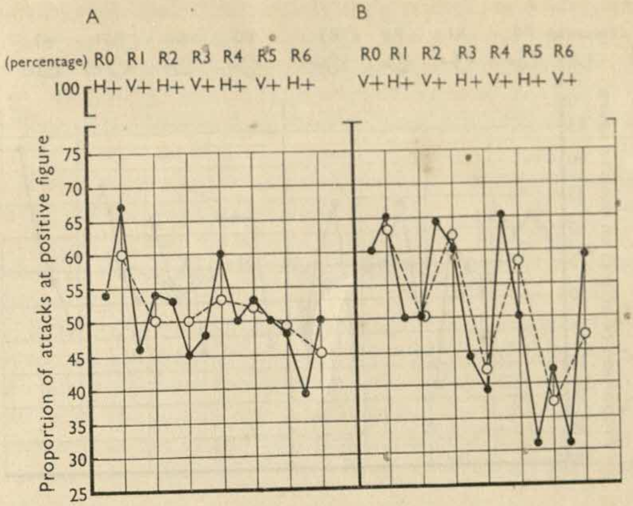
TABLE II

	Group 2					*Group 3*				
	H and V					*Square and Diamond*				
	1	2	3	4	5	6	7	8	9	10
	H+	*Percentage*	*V+*	*Percentage*	*Percentage*	□	*Percentage*	◇	*Percentage*	*Percentage*
R_0	$\frac{31}{52}$	60	$\frac{27}{43}$	63	61	$\frac{41}{72}$	56	$\frac{26}{65}$	47	56
R_1	$\frac{26}{52}$	50	$\frac{21}{42}$	50	50	$\frac{44}{88}$	50	$\frac{51}{100}$	51	51
R_2	$\frac{30}{61}$	49	$\frac{33}{53}$	62	56	$\frac{57}{102}$	56	$\frac{41}{72}$	57	55
R_3	$\frac{25}{47}$	53	$\frac{19}{45}$	42	48	$\frac{47}{90}$	52	$\frac{32}{56}$	57	54
R_4	$\frac{33}{64}$	52	$\frac{23}{40}$	58	55					
R_5	$\frac{19}{39}$	49	$\frac{14}{38}$	37	43					
R_6	$\frac{22}{49}$	45	$\frac{14}{30}$	47	46					

Group 2. Results of initial training with horizontal and vertical rectangle (R_0) and six reversals (R_{1-6}), calculated as in Table I. Columns 1 and 3 show the attacks at the positive figure, above and total attacks, below, for four animals with horizontal initially positive and four with vertical positive. Columns 2 and 4 give the percentage of attacks at the positive figure to all attacks $\frac{x}{x+y} \cdot 100$ for each of the two groups of animals separately and Column 5 for all eight animals combined. Columns 6–10 show for Group 3 similar data with eight animals for discrimination between "square" and "diamond."

In both groups there were fewer attacks during the days of the odd-numbered reversals. There was also a decline in the proportion of attacks to trials throughout the experiment. This was more marked in the group where the vertical rectangle was the original positive figure, perhaps because of the large number of shocks received during the days when the vertical figure was negative.

FIGURE 8



Group 2. Proportion of attacks at positive figure as percentage of all attacks $\left(\frac{x}{x+y} \times 100\right)$. A, when horizontal rectangle was initially positive; B, vertical initially positive. Filled circles, individual sessions, open circles complete days (two sessions).

As the level of attack on both positive and negative figures fell the animals in which the vertical rectangle was the positive figure in initial training began to show the phenomenon of increasing accuracy of response as judged by the proportion of attacks at the positive figure (Figure 8). This was not so marked as in the later stages of the reversals of black and white, but those reversals continued for a longer period of time. The phenomenon did not appear when the horizontal was the initially positive figure.

Group 3. *Repeated reversal with "square" and diamond*

A similar experiment to the last was performed using 5 cm. white squares shown either with the base parallel to the bottom of the tank ("square") or on edge ("diamond"). This discrimination is rather difficult for an octopus (Boycott and Young, 1957). Reversals were therefore given after every two days (40 trials). Food was given for each positive trial and the shocks were 6v.

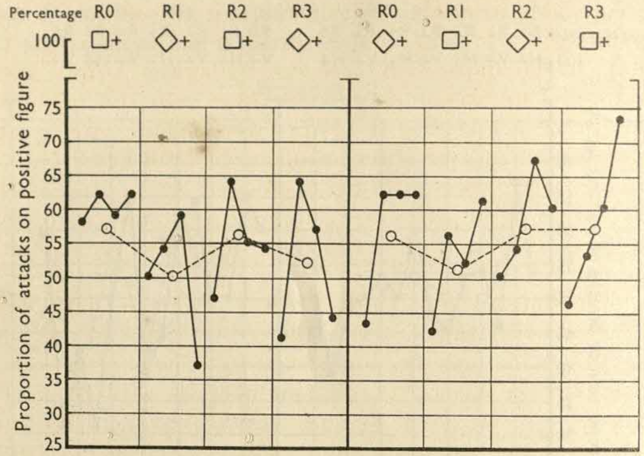
Only three reversals were completed but they show interesting confirmation of the previous experiments. The frequency of attack at either figure declined and at the same time the proportion of attacks to trials with the positive figure increased towards the end of each period of training in one direction (Table II and Figure 9). This increase was marked only in the animals in which the diamond was the initially positive figure, otherwise the two groups showed the same performance, discriminating immediately after reversal in the previously learned direction and then showing improvement at later sessions, though with some irregularities.

Group 4. *Animals without vertical lobes*

Initial learning

These nine animals had been operated 18 days previously. In all except one less than 10 per cent. of vertical lobe tissue remained. The mean amount removed was 93 per cent. (p. 208) with no serious damage to any underlying structures.

FIGURE 9



Group 3. Proportion of attacks at positive figure as percentage of all attacks $\left(\frac{x}{x+y} \times 100\right)$. A, when "square" was initially positive; B, "diamond" initially positive. Filled circles, individual sessions, open circles means for four sessions.

TABLE III

	Attacks		$\frac{x}{x+y} \cdot 100$	$\frac{x}{x+y} \cdot 100$	$\frac{x}{x+y} \cdot 100$	*Errors per cent.*		
	x +	y −		*B*+	*W*+		*B*+	*W*+
	1	2	3	4	5	6	7	8
R_0	30	40	43	49		61	51	
	31	23	57			41		
R_1	25	39	39		40	65		64
	25	36	41			62		
R_2	14	38	27	36		77	67	
	23	29	44			57		
R_3	29	24	55		53	44		47
	22	22	50			50		
R_4	22	22	50	52		50	48	
	26	22	54			46		
R_5	18	16	53		48	48		52
	16	21	43			56		
R_6	5	13	28	38		59	58	
	16	21	43			56		
R_7	14	8	64		57	43		47
	12	12	50			50		
R_8	20	24	45	47		54	52	
	10	10	50			50		

Group 4. Nine octopuses without vertical lobes. Initial learning (R_0) and eight reversals. Columns exactly as Table I. Total of 45 trials with each figure at each session.

The animals were trained with black and white circles, exactly as Group 1, that is with food for each trial with the positive figure and 11v. shock for the negative. They proved to be able to produce within the first two sessions some signs of performance that indicate representations ensuring this discrimination (Table III and Figures 1 and 3). In the first session there were 61 per cent. errors, and in the second 41 per cent. errors. The distribution of errors was quite different from that in the normal animals (p. 209). At first the errors were mainly attacks at the negative (white) figure. Later there were fewer mistakes with this figure but some failure to attack the black. In the second session there were 31/45 attacks at the positive, 23/45 at the negative figure. However, improvement was progressive throughout the second session so that at the last two trials of the day these animals made only 4/18 errors (the normals 5/20).

Individual differences were large (Figure 10). Two animals made only two mistakes over the second ten trials (a level achieved by six of the normals), one made three, and two each made four, five and six errors. The only obvious correlation of these variations with the lesions was that number 15, with 25 per cent. of intact vertical lobe tissue, was one of the best performers, but the other animal that had made only two mistakes had only 7 per cent. of vertical lobe tissue remaining (number 3).

First reversal

On reversal these animals began with nearly the maximum of errors, showing that effective representations had survived from the previous day (Figure 1). As the experiment proceeded they showed some increase in the number of attacks at the white circle, though during the afternoon this tendency gradually disappeared again. Throughout the day they only slightly reduced their tendency to attack the black circle. At the end of the day therefore, they were still showing "perverse" behaviour, attacking the positive figure less than the negative (Figure 1). Over the last ten trials of this second day there were 62 per cent. of mistakes, 44 per cent. with the white (positive) and 80 per cent. with the black (negative) figure. At the last pair of trials of the day there were 12/18 errors. Clearly no effective representations ensuring reversal of the previously established tendencies had been set up.

One animal made only three mistakes over this period, number 1 (Figure 10), but this animal had been a poor learner in the initial period. Six animals made five to seven errors, one made nine and one ten! These latter two with highly perverse behaviour (numbers 11 and 19) had both learnt well in the initial period.

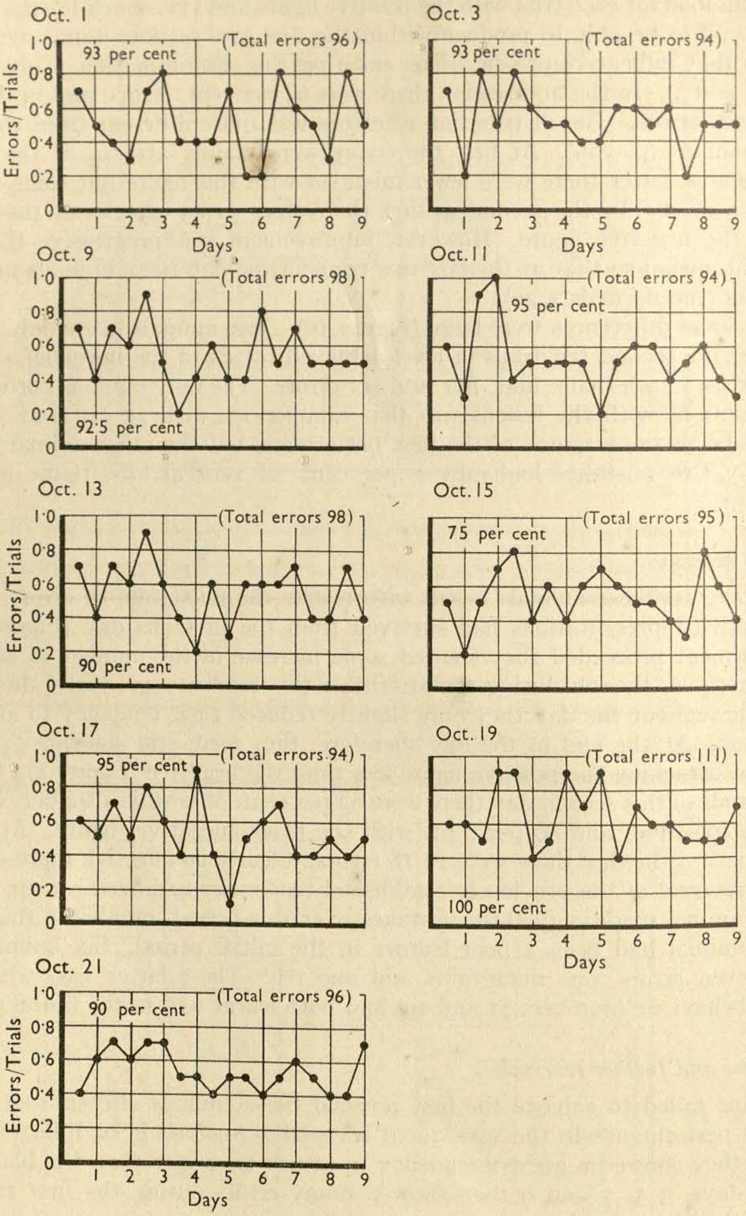
Second and further reversals

Having failed to achieve the first reversal these animals also showed no signs of correct performance in the subsequent reversals. Analysis is complicated by the fact that they showed a greater tendency to attack the white than the black circle. Thus on days, 3 5, 7 and 9 they showed many errors during the first trials, but fewer on the alternate days.

During the third day performance improved somewhat, but only to a 60 per cent. error level with both figures during the last pair of trials. During the second session of this day they made 57 per cent. errors as compared with 41 per cent. by the normals.

With the subsequent reversals there was no sign of improvement during the days, whether judged by total errors (Figure 3), errors with each colour separately (Figure 4) or errors at the last trials of each day.

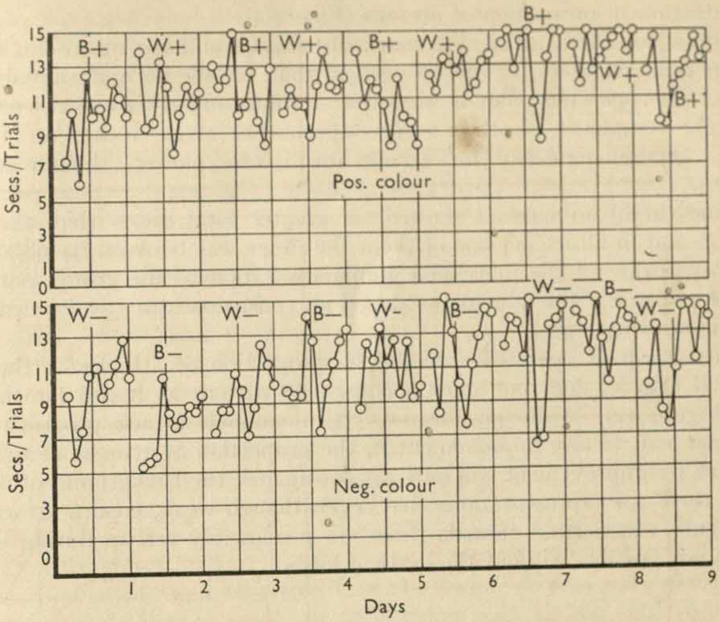
FIGURE 10



Group 4. Individual error scores of nine octopuses without vertical lobes during initial learning with a black circle positive and white negative (day 1) and eight reversals (days 2–9).

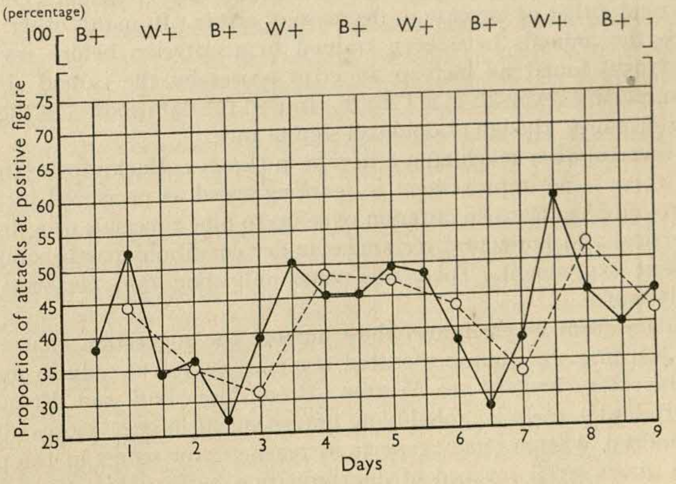
When all the reversals are plotted together therefore the pattern of performance is chaotic (Figure 5). The only general tendency was for there to be fewer mistakes with the negative figure and more with the positive towards the end of each day, in fact for the animals to come out less often to attack. Nevertheless the next

FIGURE 11



Group 4. Mean times recorded by nine animals without vertical lobes for attacks at the positive and negative colours. Conventions as for Figure 6.

FIGURE 12



Group 4. Proportion of attacks at the positive figure as percentage of all attacks $\left(\frac{x}{x+y} \times 100\right)$. Filled circles, individual sessions, open circles means for complete days (two sessions).

morning they mostly attacked both colours and there was little consistent trend to do this less often towards the end of the experiment. There was however a distinct overall reduction in proportion of attacks (Figure 4).

The times for the attacks were markedly shorter and the means for the whole group were more variable for the no vertical than for the normal animals (Figures 6 and 11). The operated animals were less cautious and their attacks once begun were usually completed, so that fewer "approaches" were recorded. There was however a general tendency for attacks to become slower, though with wide fluctuations.

The individual animals all showed far greater total errors than the normals (Figure 10), and in all except one of them the range was between 94–98/180 errors. The general feature of the individual animals was that of the group, namely irregularity of response. The animals produced no consistent strategy for dealing with these changing situations.

In this group of operated animals the proportion of attacks at the positive figure to all attacks does not show a consistent pattern as it did for the normal animals (Figure 12). Although the no vertical animals attack progressively less, this does *not* lead to any improvement in the proportion of attacks at the positive figure. Such an improvement will only appear against the background of decreasing attacks if there are representations that, even though weak, ensure attacks at the positive figure. In normal animals these are presumably set up during each day, but this is not possible without the vertical lobes.

Discussion

In terms of total error scores there was certainly no dramatic improvement by octopuses in solving these problems on successive reversals. This agrees with the experiments of Datta, Milstein and Bitterman (1960) with the crab *Gecarcinus* using somewhat similar conditions in so far as reversal of training was repeated regularly every four days (40 trials) irrespective of any criterion. They found no consistent trend either of increase or decrease of errors. In many other experiments of this type the animals have been trained to a criterion before reversal. Thus Thompson (1957) found no increase in error scores by the isopod *Armadillidium* over eight successive reversals in a T maze. Indeed the data showed a slight improvement in performance, though of doubtful significance.

Using reversal after reaching a criterion in *Octopus*, Mackintosh (1962) has also found that there is no improvement in learning speed as measured by the number of attacks required to learn to criterion over up to nine reversals of a discrimination between vertical and horizontal rectangles under conditions partly similar to those of the present experiment. There was some indication that performance became progressively worse.

These data from invertebrates thus support the suggestion that the capacity to form a "learning set" during repeated reversals is present only in mammals and perhaps birds (for chickens see Warren, Brookshire, Ball and Reynolds, 1960). In lower vertebrates there is probably no improvement in error scores after repeated reversals. Indeed Warren (1960) reports increasing error scores in the paradise fish *Macropodus* given serial reversal of discrimination habits when trained with combined spatial and brightness cues. Other experiments on fishes using various species and conditions have also failed to show progressive improvement (Adkins and Fields, 1957; Bitterman, Wodinsky and Candland, 1958). Seidman (1949) reported improving performance during repeated reversal in newts and terrapins, especially the latter, but these experiments are the only ones to report such improvement in animals other than birds and mammals, except for one individual fish *Tilapia* (Bitterman, Wodinsky and Candland, 1958).

There are however indications in the present data that even a relatively simple nervous system such as that of the octopus may have the essential elements of a mechanism for providing a strategy that will meet the situation. Together with a decreasing tendency to attack, the animals in Groups 1 and 3 showed a considerable improvement in accuracy of discrimination when judged by proportion of attacks made at the positive figure to all attacks. By this criterion the performance sometimes became strikingly good towards the ends of the days and improved as the experiment proceeded.

A result of this sort can well be imagined to occur if the representations produced by learning depend upon conditioning of cells to produce either attack or retreat (Young, 1961). As a result of repeated reversal approximately equal numbers of cells representing say "black" are conditioned to produce "attack" and "retreat." For each colour there are thus two representations, one tending to promote attack and the other to prevent it. These being approximately equal, they prevent the animal from making exploratory attacks. In the early part of each day therefore the animal shows cautious approaches but few attacks. Then, since rewards are given for every showing of the positive figure, whether or not an attack occurs, the number of cells tending to produce "attack" becomes dominant and performance at the end of the day is better.

It has previously been suggested that the supply of cells for conditioning may become exhausted (Young, 1961). It seems that most of the cells responsible for encoding "black" become conditioned during initial learning to give a positive response. Performance afterwards was always more accurate when black was the positive colour. However, there must be a mechanism remaining available to weight the response slightly as each day proceeds.

The fact that the animals without vertical lobes showed no sign at all of regular performance agrees with this hypothesis if we assume that part of the function of these lobes is to prevent the immediate effect of recent rewards and punishments from swamping the response tendencies produced by representations in the memory. The operated animals, like the normals, showed progressively fewer attacks as the series proceeded but this did not allow them to become more accurate. Being less "cautious" than the normals they were stimulated to attack the negative as often as the positive figure so that no slight balance in favour of the "correct" representation could show itself.

This hypothesis suggests that the "learning to reverse" really consists in learning not to decide at all until some slight evidence is available as to the correct direction. In the present experiment the food given at each positive trial (except at the end of the experiment with Group 1) provided the evidence once the animals were sufficiently inhibited. It would be interesting to discover whether a similar procedure would produce signs of "learning to reverse" in other species of animals that have not so far shown them. Presumably the phenomenon appears in mammals (and perhaps birds and reptiles) because the single clue provided by either one correct or one incorrect response is able to weight the system, which is kept delicately balanced by the suitable restraints imposed by the higher nervous centres.

SUMMARY

Octopuses were trained by successive presentations to discriminate between two figures by giving food for attacks at one of them and shocks for attacks at the other.

Ten trials with each figure were given during the day and the next day the rewards were reversed, food being given for attacks at the figure that previously elicited shocks and vice versa. On the third day the rewards were again reversed and so on up to nine reversals.

As judged by the proportion of errors to trials, the performance became progressively less accurate with successive reversals until it reached a random level. This was, however, largely due to a continual fall in the total number of attacks. The few attacks that were made were mainly at the positive figure. Judged by this criterion the performance became more accurate during each day and perhaps progressively more accurate on successive days.

Octopuses without vertical lobes did not show this increasing accuracy either within each day or over successive days. They did, however, show a fall in the total number of attacks.

It is concluded that progressive improvement in capacity to perform reversed discriminations under these conditions depends upon the setting up of equal representations ensuring attack at a given figure and retreat from it. These representations balance and thus prevent attack until the clue of food ensures attack at the new positive figure. In animals without vertical lobes the balance is not preserved, both figures are attacked after feeding.

REFERENCES

ADKINS, R. J., and FIELDS, P. E. (1957). Conditioning young stickhead trout to coloured lights. *Univ. Wash. Sch. Fish. Tech.*, Rep. No. 33.

BITTERMAN, M. E., WODINSKY, J., and CANDLAND, D. K. (1958). Some comparative psychology. *Amer. J. Psychol.*, **71**, 94–110.

BOYCOTT, B. B., and YOUNG, J. Z. (1957). Effects of interference with the vertical lobe on visual discriminations in *Octopus vulgaris* Lamarck. *Proc. roy. Soc., B.*, **146**, 439–59.

CRONHOLM, J. N., WARREN, J. M., and HARA, K. (1960). Distribution of training and reversal learning by cats. *J. genet. Psychol.*, **96**, 105–13.

DATTA, L. E. G., MILSTEIN, S., and BITTERMAN, M. E. (1960). Habit reversal in the crab. *J. comp. physiol. Psychol.*, **53**, 275–8.

HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.*, **56**, 51–62.

MACKINTOSH, J. (1962). An investigation of reversal learning in *Octopus vulgaris* Lamarck. *Quart. J. exp. Psychol.*, **14**, 15–22.

SEIDMAN, E. (1949). Relative ability of the newt and the terrapin to reverse a direction habit. *J. comp. Psychol.*, **42**, 320–7.

SUTHERLAND, N. S. (1960). The visual discrimination of shape by *Octopus:* squares and rectangles. *J. comp. physiol. Psychol.*, **53**, 95–103.

THOMPSON, R. (1957). Successive reversal of a position habit in an invertebrate. *Science*, **126**, 163–4.

WARREN, J. M. (1960). Reversal learning by paradise fish (*Macropodus opercularis*). *J. comp. Physiol.*, **53**, 376–8.

WARREN, J. M., BROOKSHIRE, K. H., BALL, G. G., and REYNOLDS, D. V. (1960). Reversal learning by white leghorn chicks. *J. comp. physiol. Psychol.*, **53**, 371–5.

YOUNG, J. Z. (1961). Learning and discrimination in the octopus. *Biol. Rev.*, **36**, 32–96.

YOUNG, J. Z. (1962). Reversal of learning in Octopus, and the effect of removal of the vertical lobe. *Quart. J. exp. Psychol.*, **14**, 193–205.

THE EFFECT OF AN AFTER-COMING RANDOM PATTERN ON THE PERCEPTION OF BRIEF VISUAL STIMULI

BY

M. KINSBOURNE and ELIZABETH K. WARRINGTON

From the National Hospital, Queen Square, London, W.C.1

The perception of briefly exposed visual forms is shown to be masked by an after-coming random pattern stimulus of approximately equal intensity. This effect occurs only under certain well defined conditions; it is limited by the minimum stimulus exposure time in excess of threshold which overcomes masking (critical stimulus duration) as well as by the minimum interval between presentation of the two stimuli which permits evasion of the masking action (critical interval). Over the range of stimulus duration in which masking occurs, critical interval varied with stimulus duration in such a way that the interval multiplied by the stimulus duration equals a constant. Critical stimulus duration and critical interval at threshold are shown to vary little under a variety of conditions. The effect of the random pattern stimulus is limited to the part of the visual field to which it is presented.

INTRODUCTION

It has long been known that perception of a brief visual stimulus may be masked by application of a second stimulus following hard upon the first. An early demonstration was reported by Baxt (1871), working in Helmholz's laboratory. He found that if brief exposures of single letters were followed within a short period of time by a flash of light, perception of the letters became impossible. The maximum length of the interval after which this effect could still be obtained he called the perception time–defined as "that time, after which a subsequent more intense stimulus no longer obstructs perception." He argued that perception time is constant for each category of stimulus; thus it is "10 millisec. for a letter." It represents the time needed for cerebral cortical processes to elaborate the percept; if they are disturbed during this period by the arrival of the second more intense stimulus, perception is thereby made impossible. Underlying this argument is the tacit assumption that the time interval separating the two stimuli as they impinge upon the retina is maintained when the corresponding excitations travel at the same speed between receptor and cortex. But this is not necessarily so. Bernhard (1940), by simultaneous recording of electroretinogram and electroencephalogram in man, showed that the more intense the stimulus, the sooner is electrical change detectable in both retina and cortex. The decreased latency with more intense stimulation is chiefly due to decrease in retinal delay, but to a lesser extent "post-retinal" processes are also accelerated (Monnier, 1952). Piéron (1925) regarded interference by the after-coming more intense stimulus as entirely due to the shorter latency of the latter. He took the maximum interval between the two stimuli which still permitted the occurrence of masking (*i.e.* Baxt's perception time), to be a direct measure of the diminished latency of the excitations set up by the more intense stimulus—"between a threshold excitation and a very intense excitation there is a difference of about 150 millisec., which is a measure of the amount by which the latent period may be reduced." Even Baxt (1871) had reported without comment that perception time was in fact greater the more intense he made the after-coming flash. These findings make it impossible to accept the time it takes to perceive a letter as being measured by the interval after which the letter can still be obscured by an after-coming more intense non-informational stimulus (Baxt, 1871). Only if the effect can be elicited with successive stimuli of equal intensity can Baxt's view by entertained.

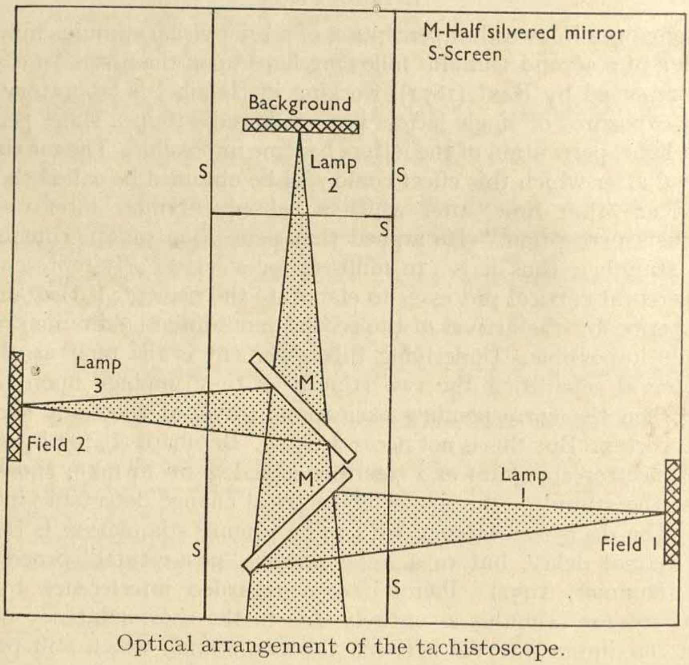
The apparent fusion of successive light stimuli of equal intensity was studied as long ago as 1875 by Exner; but this situation, which is basic to the commonly used test of "flicker fusion," is of great complexity and is no guide to the persistence time of a visual stimulus (Piéron, 1934). It does not necessarily involve the suppression of the first stimulus by intervention of the second, but merely fusion of two to appear as one. In order to produce masking not due to latency factors it is preferable to establish a situation in which a stimulus of definite and familiar contour, such as a letter, is followed by a stimulus of "non-informational" contour, both being exposed at approximately equal intensity.

The experiments described below were designed to resolve the question whether masking by an after-coming non-informational stimulus can occur in a situation uncomplicated by differences in latency; and if so, to define the limits, in terms of stimulus duration and interval between stimuli, of the occurrence of the phenomenon.

METHOD

A new Dodge type tachistoscope was used in the experiments. The optical arrangements are shown in Figure 1.

FIGURE 1



Optical arrangement of the tachistoscope.

The background and exposure field are illuminated with special fluorescent tubes having a short decay time and at the same time providing a steady light output for the duration of the exposure of the figure. (This was insured by operating the tubes with direct current). In order to eliminate any "warm up" and to provide the necessary electron emission, the filaments at each end of the tube are heated by means of a separate low-voltage winding on a transformer.

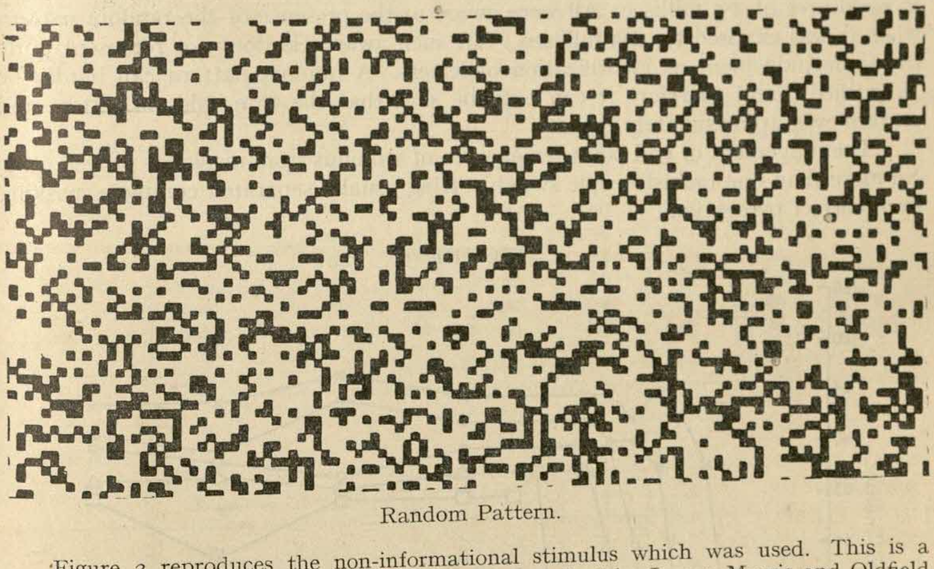
The light output of each fluorescent tube is regulated by making the tube the anode load of a triode valve, and controlling both the intensity and duration of this output by the application of a suitable switching pulse to the grid of the valve. As three tubes are used timed switching pulses are required and are provided by three identical cathode coupled monostable multivibrators. The initiation of number one multivibrator is by a push button, a timed pulse being produced for the first exposure whilst simultaneously the background illumination is first reduced, then restored. The second circuit, triggered from the first, gives an adjustable delay. The third multivibrator is then triggered from the delay circuit to the second exposure. At the same time the current in the background tube is again reduced to zero and at the end of the second exposure is restored to normal intensity.

The components used in the multivibrators circuit were calculated to give timed periods adjustable from 2 to 1600 millisec., and to ensure that results would be repeatable switched steps were chosen in preference to continuous control. It was arranged that ten switch points would give a ten fold increase in exposure period at the same time maintaining a common ratio. This ratio is approximately 1·25 and the resultant series is 2·0, 2·5, 3·2, 4·0, 5·0, 6·4, 8·0, 10·0, 12·5 and 16·0 millisec. Multiplying factors of 10 and 100 were provided by switching the timing capacitor in decade steps.

The size of the background and exposure fields is 6 × 4 in., which subtends a horizontal angle of 14° at the retina (7° either side of the fixation point). The illumination of the background and exposure fields is 2·5 foot lamberts.

The stimulus letters used were capitals, and included all those in the alphabet except I, J, O and Q. They were printed singly in black on white cards by the Letraset process*; their height was 1·2 cm. and their width varied between 1·0 cm. and 2·0 cm. The component strokes were 0·3 cm. thick.

FIGURE 2



Random Pattern.

Figure 2 reproduces the non-informational stimulus which was used. This is a section of the random pattern ("visual noise") described by Laner, Morris and Oldfield (1957)—type 80 units/sq.cm.

Six normal subjects acted as observers in these experiments. They were tested in the light adapted state.

The subject fixated binocularly a faint central mark in the tachistoscopic field. He was warned immediately before each test exposure. Using the test letters in random order, the tachistoscopic threshold for their correct perception in central vision was established, as was the threshold for the random pattern. For purposes of this work the threshold for the test letters was taken to be the shortest stimulus duration at which the subject correctly identified all of five consecutive test letters. Threshold for the random pattern was the shortest exposure duration at which the subjects were aware of its presence.

---

* Letraset Ltd., 7 Plough Place, London E.C.4.

Sets of readings were then taken for stimulus durations from threshold upwards. At each stimulus duration, various exposure durations of the after-coming random pattern were tested as follows:—

Interval was increased stepwise until the random pattern, at the exposure durations under test, no longer effectively blocked the perception of the stimulus. A single trial only was given at each "step" of interval; the first correct identification of the stimulus was taken as end point. No run of exposures was repeated more than once.

At each stimulus level runs of exposures were given at random pattern durations ranging from minimum (2 millisec.) to maximum (1600 millisec.); for each of these the minimum interval for masking of perception was established. The stimulus level was successively raised until the blocking effect could no longer be obtained under any set of conditions within the range of the tachistoscope.

EXPERIMENT I

The effect of an after-coming random pattern on the perception of single letter stimuli was systematically explored in six subjects by varying exposure duration of letter stimuli, exposure duration of random pattern and length of interval between them. In every case the end point was the minimum interval between test and random pattern stimulus at which the test stimulus was correctly identified.

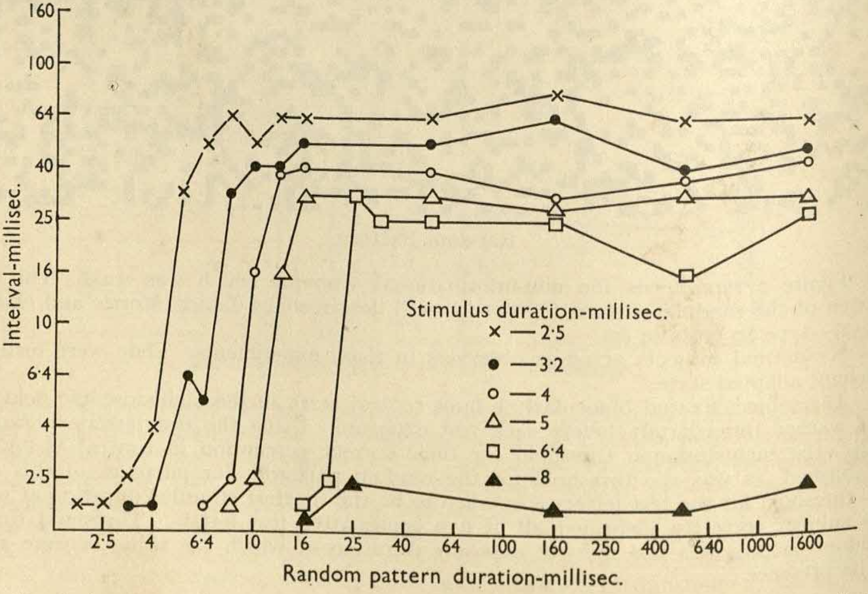
The procedure was repeated in three subjects after an interval of several days.

RESULTS

All subjects achieved complete accuracy in the identification of single letters at exposures of 2·5 millisec. All were aware of the presence of the random pattern when it was exposed for 2·5 millisec.; but such awareness does not represent quite the same judgement as identification of letters. A random pattern can hardly be "identified," and therefore it can only be said that the threshold for letters and pattern was approximately equal.

The judgement of end point (emergence of stimulus from masking) proved easy. No more than one tachistoscopic switch setting usually separated complete masking from correct perception.

FIGURE 3



Relation between random pattern duration and interval at various stimulus durations.

Table I give the readings obtained on the six experimental subjects. Figure 3 is constructed from these results to show how interval varied with duration of random pattern at various stimulus durations. Each point in the graph represents the median value of the minimum interval between test stimulus and random stimulus at which the test stimulus was correctly identified by the six subjects. The test stimulus duration was varied between 2·5 and 8·0 millisec. (above which no masking effect could be obtained). The random pattern duration was varied between 2·5 and 1600 millisec. It is apparent that the masking affect of the after-coming random pattern varies in a discontinuous way with its exposure duration. The effect reaches its maximum sharply. Further increase in exposure time of the pattern does not enhance the effect either with regard to the stimuli it can mask, or to the interval at which it can do so. For practical purposes the effect is all-or-none.

TABLE I

MINIMUM INTERVAL PERMITTING IDENTIFICATION OF TEST STIMULUS

Test stimulus	*2·5 millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	2 ..	2	2	2	2	2	2	2
	2·5 ..	2·5	2	2	2	2	2	2
	3·2 ..	5	2	2·5	3·2	2	2	2·5
	4 ..	4	2	32	4	2	2	4
	5 ..	5	32	80	50	2·5	2	32
	6·4 ..	50	32	64	50	2	3·2	50
	8 ..	80	50	80	64	2	3·2	64
	10 ..	80	40	80	50	40	2	50
	12·5 ..	64	40	64	50	64	25	64
	16 ..	64	32	80	50	64	40	64
	50 ..	100	50	64	64	40	64	64
	160 ..	80	40	80	80	80	64	80
	500 ..	80	40	100	64	64	40	64
	1600 ..	80	80	64	64	40	32	64

Test stimulus	*3·2 millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	2 ..	2	2	2	2	2	2	2
	2·5 ..	3·2	2	2	2	2	2	2
	3·2 ..	3·2	2	2	2·5	2	2	2
	4 ..	2	2	2	2	2	2	2
	5 ..	6·4	3·2	6·4	40	2	2	6·4
	6·4 ..	5	2·5	50	50	2	2	5
	8 ..	32	32	50	40	3·2	2	32
	10 ..	50	40	50	40	3·2	2	40
	12·5 ..	80	40	50	40	5·0	32	40
	16 ..	64	40	50	50	8	20	50
	50 ..	64	40	50	50	32	40	50
	160 ..	64	40	64	64	40	40	64
	500 ..	50	40	40	50	32	25	40
	1600 ..	50	50	64	50	32	32	50

Test stimulus	4·0 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	2	2	2	2	2	2	2	2
	2·5	2	2	2	2	2	2	2
	3·2	2	2	2	2	2	2	2
	4	2	2	2	2	2	2	2
	5	2	2	2	3·2	2	2	2
	6·4	2	2	4	5	2	2	2
	8	5	2	2·5	40	2	2	2·5
	10	16	10	40	40	2	2·5	16
	12·5	50	12·5	40	64	32	2	40
	16	50	32	40	50	32	2	40
	50	32	40	50	50	32	16	40
	160	25	32	40	40	32	16	32
	500	40	32	64	50	32	12·5	32
	1600	64	32	40	60	50	16	50

Test stimulus	5·0 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	5	2	2	2	2	2	2	2
	6·4	2	2	2	2	2	2	2
	8	2	2	2	2	2	2	2
	10	2·5	2	2·5	50	2·5	2	2·5
	12·5	3·2	16	50	40	2	3·2	16
	16	32	10	32	50	4·0	5·0	32
	20	40	16	32	50	5·0	5·0	32
	25	32	25	32	40	5·0	6·4	32
	32	32	16	25	40	6·4	8	25
	40	32	10	40	50	16	6·4	32
	50	32	10	32	40	25	6·4	32
	160	40	25	32	50	32	10	32
	500	40	10	32	64	25	8	32
	1600	32	25	32	50	64	5·0	32

Test stimulus	6·4 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	5	2	2	2	2	2	2	2
	6·4	2	2	2	2	2	2	2
	8	2	2	2	2·5	2	2	2
	10	2	2	2	25	2	2	2
	12·5	2	2	2	40	2	2	2
	16	2	2	2·5	50	2	2	2
	20	2	2	2·5	50	20	2	2·5
	25	40	6·4	2·5	40	20	2	20
	32	25	25	25	40	25	2·5	25
	40	25	16	25	40	20	2	25
	50	25	12·5	25	40	25	2	25
	160	25	12·5	3·2	50	20	2·5	25
	500	20	6·4	25	40	16	2·5	20
	1600	32	8	40	40	20	2	32

Test stimulus	8·0 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	5 ..	2	2	2	2	2	2	2
	6·4 ..	2	2	2	2	2	2	2
	8 ..	2	2	2	2	2	2	2
	10 ..	2	2	2	2	2	2	2
	12·5 ..	2	2	2	25	2	2	2
	16 ..	2	2	2	32	2	2	2
	20 ..	2	2	2	32	2	2	2
	25 ..	2	2	2	32	2	2	2
	32 ..	2·5	2·5	6·4	40	2	2	2·5
	40 ..	2	2	10	32	2	2	2
	50 ..	2·5	2·5	6·4	32	2	2	2·5
	160 ..	2	2	10	32	2	2	2
	500 ..	2	2	16	40	2	2	2
	1600 ..	2	2·5	8	40	2	2	2·5

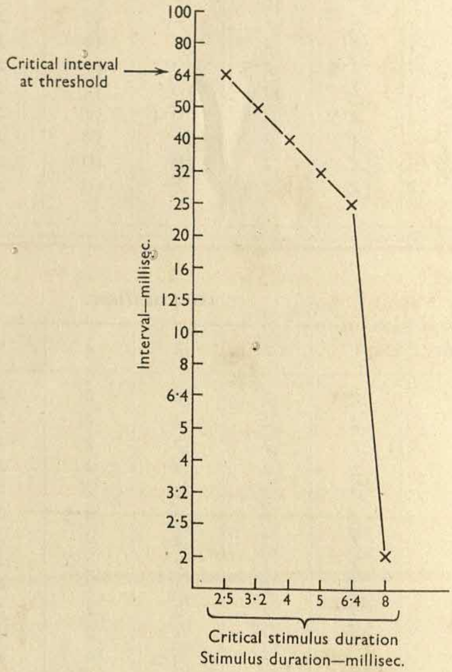
Test stimulus	10·0 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	5 ..	2	2	2	2	2	2	2
	6·4 ..	2	2	2	2	2	2	2
	8 ..	2	2	2	2	2	2	2
	10 ..	2	2	2	2	2	2	2
	12·5 ..	2	2	2	2	2	2	2
	16 ..	2	2	2	2·5	2	2	2
	20 ..	2	2	2	2	2	2	2
	25 ..	2	2	2	3·2	2	2	2
	32 ..	2	2	2	2	2	2	2
	40 ..	2	2	2	2	2	2	2
	50 ..	2	2	5	2	2	2	2
	160 ..	2	2	3·2	40	2	2	2
	500 ..	2	2	3·2	40	2	2	2
	1600 ..	2	2	3·2	32	2	2	2

Test stimulus	12·5 *millisec.*							
	Subjects	1	2	3	4	5	6	Median
Random pattern duration (millisec.)	5 ..	2	2	2	2	2	2	2
	6·4 ..	2	2	2	2	2	2	2
	8 ..	2	2	2	2	2	2	2
	10 ..	2	2	2	2	2	2	2
	12·5 ..	2	2	2	2	2	2	2
	16 ..	2	2	2	2	2	2	2
	20 ..	2	2	2	2	2	2	2
	25 ..	2	2	2	2	2	2	2
	32 ..	2	2	2	2	2	2	2
	40 ..	2	2	2	2	2	2	2
	50 ..	2	2	2	2	2	2	2
	160 ..	2	2	2	2	2	2	2
	500 ..	2	2	2	3·2	2	2	2
	1600 ..	2	2	2	2·5	2	2	2

The minimum interval between test stimulus and random stimulus permitting correct identification of test stimulus at six stimulus durations.

Figure 4 relates stimulus durations to interval. The curve was arrived at by combining the results obtained for all six subjects, at all effective random pattern exposure durations. The relationship is discontinuous. For stimulus duration at or just above threshold, increase in the duration results in decrease of minimum interval. But when a critical stimulus duration is reached, this relationship is sharply broken; the effect becomes unobtainable.

FIGURE 4



The relation between stimulus duration and interval.

The masking action of an after-coming random pattern on an informational stimulus of near equal intensity has therefore clear-cut limits under the conditions of illumination used in this study. If either stimulus duration or interval reach a critical value, the effect breaks down. The critical stimulus duration (in excess of threshold) and the critical interval (at threshold) are convenient measurements. Table II shows the critical stimulus durations and intervals obtained for the experimental subjects; it also gives these values for three of the subjects obtained when they were re-tested some days later.

It is apparent that there was little variation in the results obtained from subject to subject, and on the different occasions.

EXPERIMENT II

The same procedure was applied to the perception in correct orientation of straight lines 7·0 cm. long and 0·2 cm. thick, horizontal, vertical and diagonal, presented in random sequence. Critical stimulus durations and critical intervals are given in Table III. They differ little from those established for letter stimuli. Analysis showed that there was no significant difference in these values for horizontal and vertical, as opposed to diagonal lines—a result to be expected when a truly random pattern is used as masking stimulus.

TABLE II

CRITICAL INTERVAL AND STIMULUS DURATION

Subject	*Critical interval (at threshold)*		*Critical stimulus duration*	
	1st millisec.	*2nd millisec.*	*1st millisec.*	*2nd millisec.*
1	80	50	7·5	7·5
2	50	—	5·5	—
3	80	64	10	7·5
4	64	64	10	7·5
5	64	—	5·5	—
6	40	—	4	—

The critical interval is the minimum interval between test and after-coming random stimulus at which the test stimulus was correctly identified. The critical stimulus duration is the difference between the minimum test stimulus duration at which the masking effect was no longer obtainable and the threshold for the test stimulus as defined in "Method."

TABLE III

CRITICAL INTERVAL AND STIMULUS DURATION FOR CORRECT PERCEPTION OF LINES

Subject	*Critical stimulus duration (millisec.)*	*Critical interval (at threshold) (millisec.)*
2	5	40
3	4	50

The critical interval is the minimum interval between test and after-coming random stimulus at which the test stimulus was correctly identified. The critical stimulus duration is the difference between the minimum test stimulus duration at which the masking effect was no longer obtainable, and the threshold for the test stimulus as defined in "Method".

EXPERIMENT III

The same procedure was applied to the simultaneous perception of groups of two and three letters. Two subjects were tested by exposures given in runs as before till all letters of the group had been correctly identified. The letters were horizontally adjacent, and separated by intervals of 1 cm.

It was found that the threshold for the correct identification of all the letters in the group never exceeded that of identification of one of the group by more than one tachistoscopic step. In most cases, two and three letters were simultaneously identified at the same exposure duration as single letters. There was a similar relationship between perception of one and all when the random pattern was used to establish critical stimulus duration and critical interval (Table IV). It appears that under the present conditions simultaneous perception is influenced by the after-coming random pattern much as is single perception.

TABLE IV

CRITICAL INTERVAL AND STIMULUS DURATION FOR CORRECT PERCEPTION OF LETTER GROUPS

	Subject	*Critical stimulus duration (millisec.)*	*Critical interval (millisec.)*	*Threshold (millisec.)*
1 Letter	2	5·5	50	2·5
	3	10·0	80	2·5
2 Letters	2	5·5	64	2·5
	3	10·0	80	2·5
3 Letters	2	5·5	64	2·5
	3	13·5	100	3·2

EXPERIMENT IV

The mechanism of action of the after-coming random pattern in Experiment I and II was investigated in two subjects as follows:—

(*a*) The area of visual field to which the random pattern was applied was limited so as just to correspond to that receiving the letter stimulus. Thus a section of random pattern, 2·0 cm. × 1·2 cm., was presented centrally, the periphery of the tachistoscopic field being left clear.

Apart from this modification the procedure was identical with that of Experiment I.

The results showed that the random pattern limited to the retinal locus involved in reception of the informational stimulus is fully as effective in masking the stimulus as a random pattern covering a wider area (Table V).

TABLE V

CRITICAL INTERVAL AND STIMULUS DURATION UNDER MODIFIED CONDITIONS OF RANDOM PATTERN PRESENTATION

	Subject	*Critical stimulus duration*	*Critical interval*
Random pattern covering, but limited to locus of test stimulus	2	10·0	40
	3	7·5	50
Random pattern surrounding locus of test stimulus	2	0, at threshold	2
	3	0·7	32

The ratio of black to white within the 2·4 sq.cm. of pattern stimulus was approximately 3/7. Within the corresponding area containing the stimuli the ratio varied between 2/7 and 4/7 depending on the letter used. Thus in this experiment not only was the illumination of test and random stimulus fields the same, but it fell on visual material of comparable extension in space in terms of total area of "blackness." That the masking effect was as regularly obtained here as before, strengthens the argument that the effect has been elicited not only under conditions of equal illumination but of approximately equal effective overall intensity of the two fields.

(*b*) A blank gap was introduced into the centre of the random pattern so that though filling the rest of the field, the non-informational stimulus did not actually overlap the area of presentation of the letter stimulus.

Apart from this modification the procedure was identical with that of Experiment I.

The results showed that this form of presentation abolishes the masking effect of the after-coming random pattern (Table V). It does introduce some irregularities of threshold for perception of the test stimulus, but these are slight and variable, and not comparable to the striking and regular effect of presenting test stimulus and after-coming pattern to the same retinal area.

Discussion

The results show that under the condition of the experiments it was possible consistently to mask a test stimulus by means of an after-coming random pattern.

In its all-or-none effect in masking perception, the after-coming random pattern differs radically from the previously used non-informational stimulus, the after-coming flash. The effect of the flash varied in a continuous and complex way with stimulus duration (Baxt, 1871; Lindsley and Emmons, 1958). Piéron (1925) plausibly attributed this effect to variation of latency depending on intensity. This explanation is not applicable to a situation in which, as in the present, the stimuli are approximately equal in intensity. Piéron (1925) thought it impossible to obtain masking under such conditions—"for stimuli differing in form and of equal intensity (or nearly so) there cannot be masking; even if simultaneous, both stimuli are perceived."

Nevertheless, masking was obtained under just such conditions; and its explanation must be other than that applicable to the intense after-coming flash.

The limits of the effect of the random pattern was remarkably clear-cut; it masked stimuli of up to a critical exposure duration; when it did so, the effect was maintained until the interval between the two stimuli reached a critical value. Over the range of stimulus durations for which masking was possible (*i.e.* between threshold and critical stimulus duration) the minimum interval permitting perception varied inversely with the stimulus duration (Figure 4). An alternative statement of this relationship is: stimulus duration × interval = a constant.

Within the range of conditions tested, it was found that the critical stimulus duration regularly exceeded threshold by between 6·4 and 12·5 millisec. (median 8·0 millisec.). Such regularity must be based upon some fundamental property of visual perception. The same applies to the critical interval, which under all conditions was between 40 and 80 millisec. (median 64 millisec.). This value may be related to the minimum time intervals which can be discriminated by the visual perceptual processes; it would again appear to reflect some fundamental property of visual perception.

Baxt (1871) reported that when the groups of letter stimuli were followed by an after-coming flash, the letters were not seen simultaneously, but serially. Each additional letter required approximately an extra 10 millisec. of exposure time. Averbach and Sperling (1960) mentioned briefly that using an after-coming stimulus of "parts of letters" they arrived at similar results. Using a random pattern, this result was not obtained (Experiment III).

The reason for Baxt's finding may now never be known, as he did not report sufficient experimental details to permit his work to be repeated. Averbach and Sperling's results may be due to failure of their after-coming stimulus, not truly random, equally to mask each letter in a group. In any case the use of a random pattern does not reveal serial perception of the type postulated in "filter theory" (Broadbent, 1958). If something corresponding to a filter plays a part in visual perception, it does so at a stage in perception other than that which is affected by an after-coming random pattern.

That the masking effect is not a diffuse one is shown by the results of Experiment IV. That it is not limited to the perception of letters is shown in Experiment II; and work in progress suggests that it is widely applicable to many categories of simple forms.

Simply to state, on the lines of Baxt (1871), that the interval after which an after-coming stimulus fails to obstruct perception represents the time it takes to perceive a letter is to let hypothesis far exceed the evidence.

The concept of "perception time" is of little help in explaining the effect of an after-coming random pattern. When the cerebral cortex is required to differentiate an informational pattern from a non-informational one following it within a brief interval, that is to impose an extra stress upon the system. The task is harder, but not necessarily so because the second stimulus has "wiped out" the first (Teuber, 1960), a term which prejudges the issue. All that can be said is that the threshold exposure duration at which an informational stimulus can be differentiated from a random pattern background is higher than that at which it can be seen against a homogenous light background; and that an after-coming random stimulus effectively forms such a background when it follows within a brief period. The length of that period must indeed relate to some fundamental attribute of the neural system subserving visual perception. Perhaps this attribute is the minimum interval between two brief stimuli which permits the perceptual processes to deal with them individually, rather than as a composite stimulus which may therefore at times be unintelligible. Further experiments are required to settle this question.

We wish to thank Dr. Eliot Slater for granting facilities for this work; Mr. J. W. Whitfield for helpful comment; Mrs. E. King for help with the experimental procedures; and Mr. A. Prickett for preparing the graphs.

References

Averbach, E., and Sperling, G. (1960). Short term storage of information in vision. *Fourth London symposium on Information Theory.*

Baxt, N. (1871). Über die Zeit welche nöthig ist, damit ein Gesichtseindruck zum Bewusstsein kommt, und über die Grösse (Extension) der bewussten Wahrnehmung bei einem Gesichtseindrucke von gegebener Dauer. *Pflüg. Arch. ges. Physiol.*, **4**, 325–36.

Bernhard, C. G. (1940). Contributions to the neurophysiology of the optic pathways. *Acta. physiol. scand.*, **1**, supp. 1, 1–94.

Broadbent, D. E. (1958). *Perception and Communication.* London.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Pflüg. Arch. ges. Physiol.*, **11**, 403–32.

Laner, S., Morris, P., and Oldfield, R. C. (1957). A random pattern screen. *Quart. J. exp. Psychol.*, **9**, 105–8.

Lindsley, D. B., and Emmons, W. H. (1958). Perception time and evoked potentials. *Science*, **127**, 1061.

Monnier, M. (1952). Retinal, cortical and motor responses to photic stimulation in man. *J. Neurophysiol.*, **15**, 469–86.

Piéron, H. (1925). Recherches experimentales sur la marge de variation du temps de latence de la sensation lumineuse (par une méthode de masquage). *Année psychol.*, **26**, 1–30.

Piéron, H. (1934). L'évanouissement de la sensation lumineuse. *Année psychol.*, **35**, 1–49.

Teuber, H. L. (1960). Visual perception. In J. Field (ed.), *Handbook of Physiology, Section I, Neurophysiology, Vol. III.* Baltimore.

FURTHER STUDIES ON THE MASKING OF BRIEF VISUAL STIMULI BY A RANDOM PATTERN

BY

M. KINSBOURNE and ELIZABETH K. WARRINGTON

From the National Hospital, Queen Square, London, W.C.1

A brief visual test stimulus may be masked by a preceding random pattern of roughly equal intensity in the same way as by a succeeding random pattern. The parameters of random pattern duration, interval between stimuli, and test stimulus duration limiting the masking effect, show a regularity very similar in the two conditions. When the random pattern precedes the test stimulus, the relationship between the two may be stated as follows:—

Stimulus duration × interval = a constant.

The identical effects are found when the two stimuli are presented separately, one to each eye, both for random pattern preceding and succeeding the test stimulus. It was concluded that masking resulted through the same central interaction, whether produced by the preceding or succeeding random pattern. The interaction may be related to, and occur within the limits of, apparent simultaneity of rapidly successive visual stimuli.

INTRODUCTION

A brief visual test stimulus may be masked by an after-coming random pattern presented to the same retinal locus (Kinsbourne and Warrington, 1962). This effect occurred only within certain well defined limits: when the test stimulus duration did not exceed a critical value, at which masking is overcome (critical stimulus duration), and when the interval between the test stimulus and random pattern did not exceed a critical value, beyond which no masking effect occurred (critical interval). The exposure duration of the random pattern was of minor importance. There was a relationship between the critical interval and the test stimulus duration, which held for durations of the test stimulus up to the critical stimulus duration. This was formulated as follows:—

Stimulus duration × critical interval = a constant.

The results could not be explained by supposing a different retino-cortical latency of masked and masking stimulus, as the two stimuli were of approximately equal intensity, whereas to establish a significant difference in the latency, a vast disparity of intensity is required (Piéron, 1925; Bernhard, 1940).

An alternative suggestion was made, that the critical interval is the time which permits the perceptual process to deal with the two stimuli separately in succession, rather than simultaneously as a composite, and therefore unintelligible, stimulus.

Before such an explanation is acceptable, two further considerations must be taken into account. Can a comparable masking effect be produced when the random pattern precedes the test of stimulus? Does masking represent interaction at a cerebral or at a retinal level? Two experiments devised to answer these questions are the subject of this communication.

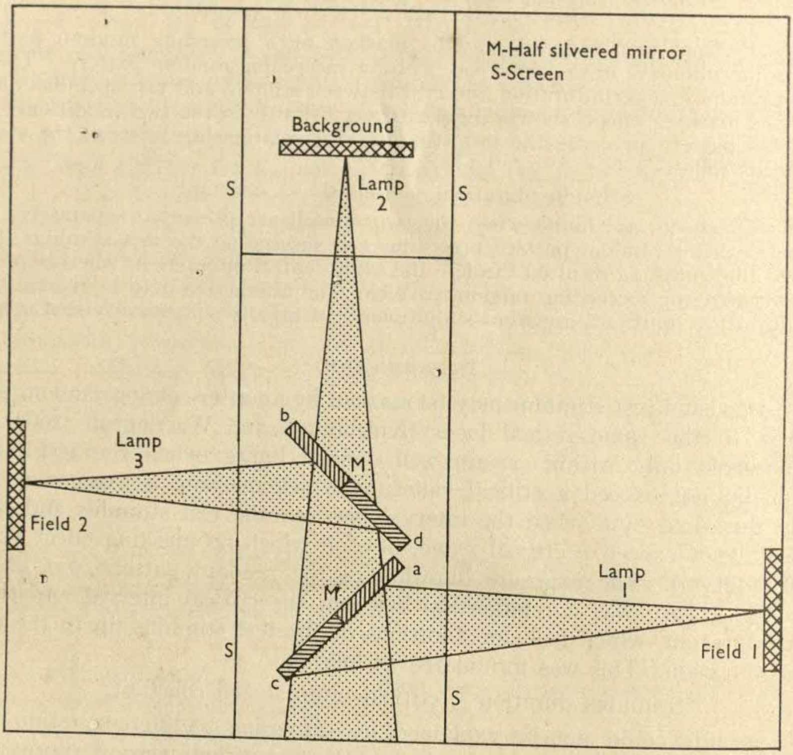
EXPERIMENT I

The aim of Experiment I was to determine the effect on a brief visual stimulus of a preceding random pattern.

Apparatus

A Dodge type tachistoscope was used. The optical arrangements are shown in Figure 1. A background field (2·5 ft. lamberts in brightness), and subtending 14 × 9° of visual angle, could be replaced by two successive test fields of equal size and brightness. In the interval between the test fields, the background field was again on view. It was possible to vary independently the exposure duration of each of the two test fields and the duration of the interval between them. For each there were thirty available exposure durations, between 2 and 1600 millisec. on a geometrical scale, with a multiplying factor of 1·25. This apparatus is described in detail by Kinsbourne and Warrington (1962).

FIGURE 1



Optical arrangement of the tachistoscope.

The test stimuli used were capital letters, and included all those in the alphabet except I, J, O and Q. They were printed singly in black on white cards by the Letraset process; their height was 1·2 cm. and their width varied between 1·0 cm. and 2·0 cm. The component strokes were 0·3 cm. thick, and exposed at a distance of 60 cm. from the eye.

The non-informational stimulus used was a section of the random pattern ("visual noise") described by Laner, Morris and Oldfield (1957)—type 80 units/sq.cm., (Fig. II in Kinsbourne and Warrington, 1962).

Seven normal subjects acted as observers in this experiment. They were tested in the light adapted state.

Procedure

The subject fixated binocularly a faint central mark in the tachistoscopic field. He was warned immediately before each test exposure. Using the test letters in random order, the tachistoscopic threshold for their correct perception in central vision was established. For purposes of this work, the threshold for the test letters was taken to be the shortest stimulus duration at which the subject correctly identified all of five consecutive test letters.

Sets of readings were then taken for stimulus durations from threshold upwards. At each stimulus duration, various exposure durations of the preceding random pattern were tested as follows:—

Interval was increased stepwise from minimum (2 millisec.) until the random pattern, at the exposure durations under test, no longer masked the perception of the stimulus. A single trial only was given at each "step" (switch setting) of interval; the first correct identification of the stimulus letter was taken as end point. No run of exposures was repeated more than once.

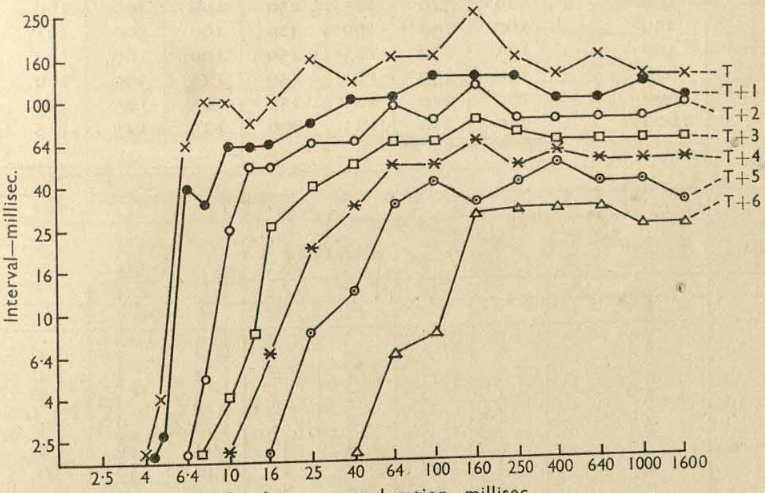
At each stimulus level runs of exposures were given at random pattern durations ranging from minimum (2 millisec.) to maximum (1600 millisec.); for each of these the minimum interval at which correct perception just became possible was established. The test stimulus duration was then successively increased and the above procedure was repeated at each test stimulus duration until the masking effect could no longer be obtained at any value of random pattern duration and interval within the range of the tachistoscope.

This procedure was identical with that used by Kinsbourne and Warrington (1962), except that the random pattern preceded the test stimulus, instead of following it.

RESULTS

The minimum interval necessary for the correct recognition of the test stimulus at each duration of random pattern, and for each duration of test stimulus (up to the duration at which masking is no longer obtained), is given in Table I. The values obtained from each subject are recorded separately, as is the median value for all subjects under any given condition. Figure 2 illustrates the minimum interval necessary for the correct recognition of the test stimulus under varying conditions of random pattern duration and test stimulus duration, using the median value for all the subjects. Thus each point on the graph represents the median of seven values (one from each subject) of the minimum interval necessary for correct recognition of the test stimulus under the given condition. The form of the family of curves is very similar to that obtained in the previous communication, in which the masking

FIGURE 2



Relation between random pattern duration and interval at various stimulus durations. (Random pattern preceding test stimulus).

T = Threshold stimulus duration.

No's 1—7 represent serial switch settings with a multiplying factor of 1·25.

effect of an after-coming random pattern on a brief visual stimulus was established (Kinsbourne and Warrington, 1962). These findings show that a preceding random pattern is as effective in masking a brief visual stimulus as is an after-coming random pattern.

TABLE I

Minimum Interval Permitting Identification of Test Stimulus

Test stimulus	*At Threshold*								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5	2	2	2	2	2	2	2	2
	3·2	2	2	2	2	2	2	2	2
	4·0	2·5	2	2	10	2	3·2	2	2
	5·0	4	2	2	64	2	64	4	4
	6·4	20	64	50	80	64	125	32	64
	8·0	160	100	64	100	100	100	100	160
	10·0	125	80	64	125	64	125	100	100
	12·5	100	160	80	100	80	80	80	80
	16·0	100	250	100	200	64	100	125	100
	25·0	160	320	80	125	100	100	250	160
	40·0	100	250	100	400	125	250	100	125
	64·0	250	250	125	640	125	160	160	160
	100	160	250	160	320	125	160	250	160
	160	320	320	125	250	100	400	160	250
	250	160	250	160	320	100	500	125	160
	400	125	250	125	250	100	160	125	125
	640	160	250	160	250	125	500	160	160
	1000	160	320	125	250	125	125	125	125
	1600	250	250	125	1000	125	125	125	125

Test stimulus	*Threshold* +1								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5	2	2	2	2	2	2	2	2
	3·2	2	2	2	2	2	2	2	2
	4·0	2	2	2	2	3·2	3·2	2	2
	5·0	2·5	2	2	2	3·2	6·4	4	2·5
	6·4	2·5	50	2	40	64	40	2	40
	8·0	32	50	2	25	40	40	10	32
	10·0	50	64	50	80	64	64	5	64
	12·5	64	80	64	125	50	50	64	64
	16·0	80	64	50	100	64	50	50	64
	25·0	100	125	80	64	64	50	160	80
	40·0	125	125	80	125	80	80	100	100
	64·0	160	125	80	80	64	100	100	100
	100	160	100	100	125	80	125	125	125
	160	160	160	125	160	100	100	80	125
	250	125	125	100	125	80	200	160	125
	400	100	125	125	100	100	100	125	100
	640	100	125	80	125	100	100	100	100
	1000	160	320	80	160	100	125	100	125
	1600	100	100	80	125	100	125	80	100

Test stimulus	*Threshold* +2								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5	2	2	2	2	2	2	2	2
	3·2	2	2	2	2	2	2	2	2
	4·0	2	2	2	2	2	2	2	2
	5·0	2·5	2	2	2	2	2	2	2
	6·4	2·5	2	3·2	2	2	2	2	2
	8·0	2·5	2	6·4	5	5	5	2	5
	10·0	20	32	40	5	50	25	2	25
	12·5	50	50	50	64	50	32	4	50
	16·0	50	64	64	50	40	40	10	50
	25·0	64	64	64	64	50	50	40	64
	40·0	125	80	64	80	64	50	64	64
	64·0	100	100	80	125	50	64	100	100
	100	80	125	64	100	64	80	80	80
	160	125	125	80	125	80	125	100	125
	250	80	100	64	100	80	80	64	80
	400	80	100	80	100	80	80	64	80
	640	80	100	64	100	100	80	80	80
	1000	80	100	80	80	80	125	80	80
	1600	64	100	100	80	100	100	80	100

Test stimulus	*Threshold* +3								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5	2	2	2	2	2	2	2	2
	3·2	2	2	2	2	2	2	2	2
	4·0	2	2	2	2	2	2	2	2
	5·0	2	2	2	2	2	2	2	2
	6·4	2	2	20	2	2	16	2	2
	8·0	2	2	20	2	2	8	2	2
	10·0	4	2	32	16	2	10	4	4
	12·5	8	2	32	32	32	16	2	8
	16·0	25	25	40	64	25	25	2	25
	25·0	80	40	50	64	40	37	20	40
	40·0	40	64	50	64	40	50	50	50
	64·0	80	50	50	100	64	64	64	64
	100	80	80	50	64	50	50	80	64
	160	100	80	50	100	64	50	80	80
	250	80	80	64	80	64	80	64	80
	400	80	80	64	64	64	64	80	64
	640	80	64	64	80	64	64	50	64
	1000	64	50	64	80	64	64	64	64
	1600	64	64	50	64	80	80	64	64

Test stimulus	*Threshold* +4								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5 ..	2	2	2	2	2	2	2	2
	3·2 ..	2	2	2	2	2	2	2	2
	4·0 ..	2	2	2	2	2	2	2	2
	5·0 ..	2	2	2	2	2	2	2	2
	6·4 ..	2	2	2	2	2	2	2	2
	8·0 ..	2	2	2	2	2	2	2	2
	10·0 ..	2	2	2	40	2	2	2	2
	12·5 ..	2	2	2	40	2	2	2	2
	16·0 ..	10	8	6·4	40	2	2·5	3·2	6·4
	25·0 ..	16	32	10	40	20	50	2	20
	40·0 ..	50	32	32	50	32	40	16	32
	64·0 ..	80	50	40	80	50	40	32	50
	100 ..	50	40	40	100	50	50	80	50
	160 ..	64	80	50	64	50	50	64	64
	250 ..	64	64	32	100	50	40	50	50
	400 ..	64	64	40	80	50	64	50	64
	640 ..	64	64	32	64	50	40	50	50
	1000 ..	64	50	40	64	50	64	40	50
	1600 ..	40	64	40	80	50	50	50	50

Test stimulus	*Threshold* +5								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	2·5 ..	2	2	2	2	2	2	2	2
	3·2 ..	2	2	2	2	2	2	2	2
	4·0 ..	2	2	2	2	2	2	2	2
	5·0 ..	2	2	2	2	2	2	2	2
	6·4 ..	2	2	2	2	2	2	2	2
	8·0 ..	2	2	2	2	2	2	2	2
	10·0 ..	2	2	2	2	2	2	2	2
	12·5 ..	2	2	2	2	2	2	2	2
	16·0 ..	16	2	2	5	2	2	2	2
	25·0 ..	20	2	2	8	25	2	10	8
	40·0 ..	40	2	12·5	64	3·2	25	4	12·5
	64·0 ..	50	25	3·2	64	2·5	50	32	32
	100 ..	40	64	25	80	40	50	32	40
	160 ..	32	50	32	50	32	64	25	32
	250 ..	40	40	32	40	32	50	40	40
	400 ..	50	50	50	40	32	50	40	50
	640 ..	50	50	25	40	40	50	40	40
	1000 ..	40	40	32	50	40	40	40	40
	1600 ..	32	32	32	64	32	40	32	32

Test stimulus	*Threshold* +6								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	25	3·2	2	2	2	2	2	2	2
	40	16	12·5	2	2	2	2·5	2	2
	64	25	2	2	64	6·4	5·0	2	6·4
	100	25	2	2	32	40	8	3·2	8
	160	32	5	2	40	32	32	6·4	32
	250	32	2	2	50	32	32	25	32
	400	40	2	2	64	32	32	40	32
	640	32	8	2	64	32	40	3·2	32
	1000	25	2	4	40	32	25	2	25
	1600	25	2	2	40	32	32	3·2	25

Test stimulus	*Threshold* +7								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	25	2	2	2	2	2	2	2	2
	40	2	2	2	2	2	2	2	2
	64	3·2	2	2	25	12·5	2	2	2
	100	4·0	2	2	40	2	5	2	2
	160	4·0	2	2	40	3·2	3·2	8	3·2
	250	2·5	2	2	40	2	2	32	2
	400	2·5	2	2	25	3·2	2·5	20	2·5
	640	3·2	2	2	40	25	2	2	2
	1000	2·5	2	2	32	25	2	2	2
	1600	3·2	2	2	40	25	2·5	2	2·5

Test stimulus	*Threshold* +8								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	25	2	2	2	2	2	2	2	2
	40	2	2	2	2	2	2	2	2
	64	2	2	2	2	2	2	2	2
	100	2	2	2	2	2	2	2	2
	160	2	2	2	2	2	2	2	2
	250	2	2	2	32	2	2	2	2
	400	2	2	2	32	2	2	2	2
	640	2	2	2	32	2	2	2	2
	1000	2	2	2	20	2	2	2	2
	1600	2	2	2	25	2	2	2	2

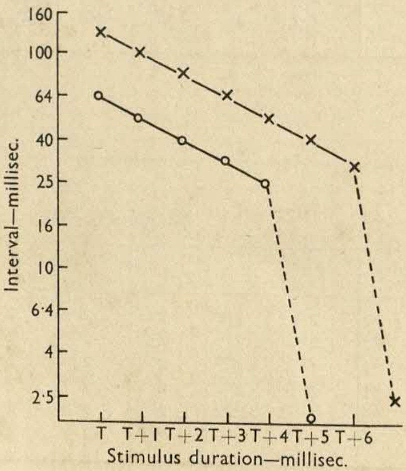
Test stimulus	*Threshold* +9								
	Subjects	1	2	3	4	5	6	7	Median
Random pattern duration (millisec.)	25 ..	2	2	2	2	2	2	2	2
	40 ..	2	2	2	2	2	2	2	2
	64 ..	2	2	2	2	2	2	2	2
	100 ..	2	2	2	2	2	2	2	2
	160 ..	2	2	2	2	2	2	2	2
	250 ..	2	2	2	2	2	2	2	2
	400 ..	2	2	3·2	2	2	2	2	2
	640 ..	2	2	2	2	2	2	2	2
	1000 ..	2	2	2	2	2	2	2	2
	1600 ..	2	2	2	2·5	2	2	2	2

The results are expressed in millisec. of interval.

The threshold for the seven subjects ranged from 2·5 to 4·0 millisec. Test stimulus duration was progressively raised from threshold by steps involving a multiplying factor of 1·25.

The relationship between the test stimulus duration and the minimum interval necessary for correct recognition of the test stimulus (critical interval) is illustrated in Figure 3. The critical interval for each test stimulus strength was calculated by taking the median value of the individual critical intervals obtained at random pattern durations of 200, 400, 640, 1000, 1600 millisec. (at these random pattern durations the masking effect has reached its maximum—Figure 2). Figure 3 presents the combined results of the seven subjects; each point on the curve is the critical interval at a given test stimulus strength. For comparison, the relationship between critical interval and test stimulus strength obtained when the test stimulus preceded the random pattern (Kinsbourne and Warrington, 1962) is also included.

FIGURE 3



The relation between stimulus duration and interval.

× Random pattern preceding test stimulus.

○ Random pattern following test stimulus.

The two curves are seen to be identical in slope. There are two points of difference; firstly, the absolute value of the critical intervals is greater at any given stimulus strength when the random pattern precedes the test stimulus; secondly, under these conditions the masking effect is still obtained for test stimuli of slightly greater duration. In these respects the masking effect of a preceding random pattern is somewhat greater than that when the random pattern follows the test stimulus.

EXPERIMENT II

The aim of Experiment II was to determine the effect of dichopic presentation of a brief test stimulus, and a preceding or succeeding random pattern.

Apparatus

The optical system of the tachistoscope described in Experiment I was adapted for dichopic presentation of the test stimulus and the random pattern (Figure 1). The mirrors could be so arranged that the stimulus from the first test field could be presented to either left or right eye alone (mirror in positions *a* and *c* respectively). Similarly, the stimulus from the second test field could be presented to either eye alone, the left or the right (positions *b* and *d* respectively), whichever was not receiving the stimulus from the first test field.

The test stimuli and random pattern were the same as those used in Experiment I. Three normal subjects acted as observers in this experiment.

Procedure

The parameters of masking by a preceding and by a succeeding random pattern were established. The random pattern duration was kept constant at 500 millisec. At each stimulus duration, from threshold value to that duration at which no masking was obtained, the critical interval for the correct recognition of the test stimulus was recorded, both when the random pattern preceded and when it succeeded the test stimulus. Three readings were taken under each of these conditions, and the median value was taken as the critical interval in each case.

In two subjects the right eye received the first stimulus when the random pattern preceded the test stimulus, and in two subjects the right eye received the first stimulus when the random pattern followed the test stimulus.

TABLE II

MONOCULAR PRESENTATION OF TEST STIMULI AND RANDOM PATTERN

		Random pattern preceding				*Random pattern succeeding*			
	Subject	1	2	3	Median	1	2	3	Median
		Interval (millisec.)				Interval (millisec.)			
Test stimulus duration (millisec.)	T	160	125	250	160	100	80	80	80
	T +1 ..	100	100	160	100	80	80	80	80
	T +2 ..	80	80	80	80	50	64	40	50
	T +3 ..	64	80	80	80	32	40	40	40
	T +4 ..	50	64	64	64	25	40	12·5	25
	T +5 ..	32	50	3·2	32	12·5	16	12·5	12·5
	T +6 ..	2·5	40	2	2·5	12·5	10	2	10
	T +7 ..	2	25	2	2	2	2	2	2
	T +8 ..	2	2	2	2	2	2	2	2

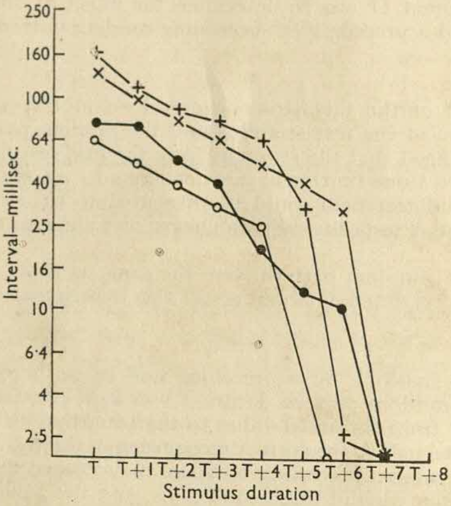
Random pattern duration was 500 millisec. in every instance.

The threshold for the three subjects ranged from 3·2 to 4·0 millisec. Test stimulus duration was progressively raised from threshold by steps involving a multiplying factor of 1·25.

RESULTS

Table II presents the critical intervals at each test stimulus duration, under conditions of separate monocular presentation. The results for each subject are given separately and combined by taking the median value. These results are shown in Figure 4. The resulting curves are seen to be very similar to those obtained with binocular presentation, both in slope and in terms of absolute values.

FIGURE 4



The relation between stimulus duration and interval.

Random pattern preceding { Binocular presentation ×
Monocular separate presentation +

Random pattern following { Binocular presentation ○
Monocular separate presentation ●

Separate monocular presentation of test stimulus and random pattern.

DISCUSSION

The masking effect occurs in the same way when test stimulus and random pattern are presented separately, one to each eye, as under binocular conditions (Kinsbourne and Warrington, 1962, and Exp. I), both when the random pattern precedes and when it succeeds the test stimulus. Evidently the effect does not depend upon interaction at the retinal level. The mechanism of masking must therefore be considered in the light of knowledge of central perceptual processes.

The results of Experiment I show that masking of a test stimulus by a preceding random pattern can as readily be obtained as from a succeeding random pattern. Moreover, although in absolute values of stimulus duration and interval limiting the effect the range within which masking occurs is greater when the random pattern is presented first, the characteristics of the interaction and its limits are so similar as to leave no doubt that both forms of masking are due to the same central mechanism. In particular, the relationship between test stimulus duration and critical interval is here, as before:—

$$\text{Stimulus duration} \times \text{critical interval} = \text{a constant.}$$

Baxt (1871) considered that masking could only occur if the masking stimulus followed the test stimulus; that is, when it arrives "centrally" during the "perception time" of the test stimulus, *i.e.*, during the time it takes for the test stimulus to be perceived. This does not suffice to explain the present findings, as the masking stimulus, although preceding the test stimulus, and thus arriving centrally outside its "perception time," nevertheless masks its perception. The results are more consistent with the consideration touched upon in the previous communication (Kinsbourne and Warrington, 1962), that there may exist a limitation of the visual perceptual process as regards the temporal succession of stimuli at, or just above, threshold. The effect of such a limitation would be that when such a stimulus is presented very soon before or after another, the two are treated as "simultaneous," and are recorded as a composite image, without temporal differentiation. When this composite image includes a random pattern of the kind used, it may prove too complex for retrospective analysis. It would then be impossible to pick out and correctly report the informational (test) stimulus.

If this is so, the limits of the subjective sensation of "simultaneity" might be found to coincide with those of the masking effect under comparable conditions. Furthermore, it might prove that other phenomena, also depending on the presence or absence of apparent simultaneity, are limited in their occurrence by parameters corresponding to those which limit the masking effect.

The experimental test of these predictions will be reported in another communication.

We wish to thank Dr. Eliot Slater for granting facilities for this work; Mrs. E. King for help with the experimental procedures; and the Department of Medical Illustration, The Hospital for Sick Children, Great Ormond Street, for preparing the graphs.

REFERENCES

BAXT, N. (1871). Über die Zeit welche nötig ist damit ein Gesichtseindruck zum Bewusstsein kommt, und über die Grösse (Extension) der bewussten Wahrnehmung bei einem Gesichtseindruck von gegebener Dauer. *Pflüg. Arch. ges. Physiol.*, **4**, 325–36.

BERNHARD, C. G. (1940). Contributions to the neurophysiology of the optic pathways. *Act. Physiol. scand.*, **1**, supp. 1.

KINSBOURNE, M., and WARRINGTON, E. K. (1962). The effect of an after-coming random pattern on the perception of brief visual stimuli. *Quart. J. exp. Psychol.* **14**, 223–34.

LANER, S., MORRIS, P., and OLDFIELD, R. C. (1957). A random pattern screen. *Quart. J. exp. Psychol.*, **9**, 105–8.

PIÉRON, H. (1925). Recherches expérimentales sur la marge de variation du temps de latence de la sensation lumineuse par une methode de masquage. *Année psychol.*, **26**, 1–49.

SHORTER ARTICLES AND NOTES

ELIMINATIVE AND ENUMERATIVE BEHAVIOUR IN A CONCEPTUAL TASK

BY

N. E. WETHERICK

From the M.R.C. Unit for Research on Occupational Aspects of Ageing, Liverpool

P. C. Wason's paper "On the Failure to Eliminate Hypotheses in a Conceptual Task" is criticized on the grounds that the task set is in important respects untypical of problem solving situations in general. It is suggested that few or no subjects are "Eliminators" in Wason's sense and that although many are "Enumerators" in his sense, enumerative behaviour is a function of the situation. An experiment is reported in which it appears that modification to the situation can sharply reduce the incidence of enumerative behaviour and increase the subject's chance of success at an early stage. The frequency with which hypotheses are directly eliminated is shown to be very low in all groups.

In a recent paper in this *Journal*, P. C. Wason discusses what he calls "The failure to eliminate hypotheses in a conceptual task." He presented to his subjects a group of three numbers "2 4 6" telling them that the group exemplified a rule that he had in mind (the rule was "any three numbers in ascending order"). The subject was required to find the rule by proposing further groups of three numbers with reasons and being told whether his group conformed to the rule or not. The subject could say what he thought the rule was at any time but the instructions emphasized that he was not to do so till he was quite sure. He found that only a minority of his subjects got the correct rule at the first attempt and interpreted this to mean that the majority were either unable or unwilling to test their hypotheses. He supported this conclusion by demonstrating significant differences between the group successful at the first attempt and the rest on two indices; an "Eliminative/Enumerative index" and a "Negative instance index." The former was calculated as a ratio between the number of subsequent instances incompatible with each reason proposed and the number compatible, summed over all the reasons proposed. The latter was a simple ratio of negative to positive instances proposed.

Wason's experiment can be criticised on two grounds. The subject's task is misleading in so far as customary methods of testing hypotheses cannot in this situation lead to a hypothesis being eliminated. Further, there is no evidence that any subject actively eliminated hypotheses in this situation; the two indices measure the extent to which a variety of hypotheses are considered instead of only one, not the extent to which hypotheses are eliminated.

In Wason's experiment the rule may be regarded as the "Truth," the "Law of Nature" which the subject regarded as a scientific investigator is trying to discover. We will call instances conforming to the rule "Positive" and instances not conforming "Negative." Instances may at the same time either conform or not conform to whatever hypothesis the subject has in mind as to the "True law of nature" in the situation. Four cases may arise; conforming-positive, conforming-negative, not conforming-positive and not conforming-negative. Wason did not ask his subjects to state whether they thought their instances would be positive or negative when they advanced them but this information is necessary if we are to state positively that a subject has knowingly eliminated a hypothesis. For example, if a subject proposes "3 5 7—odd numbers" and is told that this instance is positive he has logically eliminated the hypotheses "Even numbers ascending" and "A number, twice the number, three times the number." But neither of these hypotheses may have crossed his mind. If he thought that "3 5 7" would be positive he cannot be known by the experimenter to have eliminated any hypothesis; if, however, he thought it would be negative he has clearly tested and eliminated the hypothesis "Even numbers ascending" even though he may remain unaware that "A number, twice the number, three times the number" has also been tested and eliminated. A subject who considers a number of hypotheses generates a high eliminative index but he cannot be known by Wason to have eliminated them. Negative instances are not an indication of eliminative behaviour; a negative instance that the subject expects to be negative is enumerative in Wason's sense.

Only two of the four cases that may arise eliminate the hypothesis under consideration; the conforming-negative and the not conforming-positive. The initial instance Wason used embodies a number of obvious relationships of the general form "Ascending order with some further restriction." Any subject considering one of these hypotheses (and this includes the majority) would be at a disadvantage because any instance conforming to the subject's hypothesis would be positive *a fortiori*. He could only eliminate his hypothesis by proposing an instance which did not conform on the off chance that it might turn out positive. This appears to be what Wason means by "Testing a hypothesis" but it is not a characteristic way of proceeding in ordinary life situations. Wason might reply that this kind of test is used in scientific procedure (control groups) and that his subjects (university students) might reasonably be expected to use it where necessary. The question then arises whether the situation is one that calls for the use of a control procedure. The hallmark of such a situation is that the class of interest cannot be unequivocally distinguished from the contrast class. Where the distinguishing characteristic is a difference that is large compared with the normal deviation of the variable from its mean value the need for controls is not so obvious. In Wason's situation the difference is quite unequivocal—given by the experimenter. On ordinary scientific criteria the need for a control procedure does not arise.

The distinction shown up by Wason's indices is between subjects who having confirmed the sufficiency of one hypothesis go on to propose it as the rule (enumerators) and others who are more cautious and consider a number of hypotheses, though without necessarily eliminating any of them. Enumerators could only be right first time if they happened to pick on the correct hypothesis at once—an unlikely contingency. Wason's paper might better be entitled "On the need to propose solutions to a conceptual task." Even this need is, however, a function of the situation. There are a number of obvious possible solutions to Wason's problem. If the task is made a little more "difficult" by making possible solutions less obvious the incidence of enumerative behaviour can be reduced, as appears from the following experiment.

PROCEDURE

The procedure differed from Wason's only in the following respects:—

(1) The subject was asked to state whether he thought his instance would conform to the unknown rule when he advanced it. It was, therefore, possible to judge whether he was knowingly eliminating hypotheses.

(2) The subject wrote his groups of numbers on slips of paper which were filed under "Conforms" or "Does not conform," thus keeping positive and negative instances separate.

(3) All the recording was done by the experimenter.

(4) Subjects were divided into four experimental groups. Group A worked under the same conditions as Wason's subjects (except for the modifications in (1) to (3) above). Group B began with a negative instance "4 2 6." Groups C and D did the task as a categorization problem. Instances in one of two categories conformed to a rule but the subject was not told which. Group C began with a positive instance and Group D with a negative.

Subjects

Forty subjects were used (20 M, 20 F), taken from an introductory course in psychology. They were allocated to experimental groups randomly within sexes.

Results

Table I shows the number of subjects successful at the first, second or a later attempt. Groups B, C and D appear to have a clear advantage over Group A, though the difference does not quite reach statistical significance. Different experimental conditions do not reduce the number of subjects failing altogether, but subjects who succeed do so at an earlier stage. Six out of 10 subjects in Group A began by using enumerative procedures (on the strictest possible definition) but only four out of 30 in the remaining groups (2—B, 2—D). This difference is significant ($p = 0{\cdot}007$ Fisher exact probability, one tail). It would appear that successful subjects in Group A may have needed more attempts than those in Groups B, C or D because their experimental conditions encouraged the use of enumerative procedures.

TABLE I

Number of Subjects Finding the Correct Rule at the First or a Later Attempt (or Failing)

Rule correct	*A*	*B*	*C*	*D*	*All groups*
First	1	4	5	3	13
Second	2	3	1	2	8
Third	1	1	1	2	5
Fourth or later ..	3	1	–	–	4
Failed	3	1	3	3	10

The instances proposed were classified as either (i) Supporting the subject's current hypothesis (conforming-positive or not conforming-negative); (ii) Refuting the hypothesis (conforming-negative or not conforming-positive); or (iii) Non-informative (where the subject is not following any rational procedure). Table II shows the average number of each type of instance proposed by subjects successful at different stages or failing altogether. The average usage of type (ii) is very low, probably because whenever one of the obvious possible hypotheses was under consideration the conforming-negative case could not arise. Subjects who propose several rules before getting the correct one do not use many more instances than subjects who are right first time. Experimental conditions did not affect the number and type of instances proposed to any extent.

Groups B, C and D had what might seem to be a more difficult task; either they did not know whether the initial group conformed to the rule or not (C and D) or

TABLE II

Average Number and Type of Instances Proposed by Subjects Successful at Different Stages or Failing

	First attempt	*Second attempt*	*Third attempt*	*Fourth or later attempt*	*Failed*
Type (i) ..	5·76	4·87	6·60	7·00	6·70
Type (ii) ..	1·46	1·50	1·20	0·50	1·10
Type (iii) ..	1·93	1·63	3·20	4·00	11·40
Total	9·15	8·00	11·00	11·50	19·20

they knew that it did not (B). But this extra difficulty had the effect of improving their performance. It seems fairly clear that this improvement is related to the significant reduction in the incidence of enumerative behaviour in these groups. If there are obvious possible solutions to a problem many subjects (perhaps the majority) may feel bound to propose them. It should, however, be borne in mind that this experimental situation is not sufficiently parallel either to ordinary life problem solving or to scientific investigation to make it worthwhile to generalise on the basis of the result obtained.

I am indebted to Dr. Alastair Heron and my colleagues at the Liverpool Unit for Research on Occupational Aspects of Ageing and to Mr. D. B. Bromley of the Psychology Department, Liverpool University, for assistance and advice at all stages in the preparation of this manuscript. Also to Mrs. B. G. Brechling for assistance in obtaining subjects.

Reference

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quart. J. exp. Psychol.*, **12**, 129–40.

REPLY TO WETHERICK

BY

P. C. WASON

From the Department of Psychology, University College, London

Wetherick makes two criticisms of Wason's (1960) experiment: (1) that the task is "untypical of problem solving situations in general," and (2) that there is "no evidence that any subject actively eliminated hypotheses."

(1) Wason's task was not intended to be typical of problem solving situations in general. His aim was to investigate the acquisition of evidence in inductive reasoning. The task used was deliberately biased so that several plausible hypotheses about an unknown rule could be supported by citing instances which confirmed them, or refuted by citing instances which disconfirmed them. None of these hypotheses could be decisively confirmed, but each could be decisively eliminated by (*a*) citing an instance at variance with it, and (*b*) then observing that this instance conformed to the unknown rule. The point of the experiment was not *whether* the subjects eventually discovered this rule (which was necessarily very general), but *how* they set about trying to discover it. They could either assume that merely sufficient hypotheses were the rule (enumerative reasoning), or try to eliminate the hypotheses which occurred to them (eliminative reasoning).

It follows that Wetherick's assertion that, "the subjects' task is misleading in so far as customary methods of testing hypotheses cannot in this situation lead to a hypothesis being eliminated," is incorrect. The experimental situation is formally equivalent to the investigation of an event for which several superficial explanations are possible. An explanation can be ruled out by finding (or producing) an occurrence of the event for which that explanation cannot hold. Basically, the problem is this: under what circumstances is a possible explanation accepted as the only explanation?

(2) The distinction between enumerative and eliminative reasoning lies in the extent to which a subject's instances are at variance with his hypotheses, regardless of whether the instances conform, or not, to the unknown rule. Enumerative reasoning is clearly demonstrated by the announcement of untested rules, and by instances which simply confirm them. Eliminative reasoning is demonstrated by instances at variance with hypotheses, and by the absence of announcements of untested rules. This criterion is less satisfactory than that for enumerative reasoning because it can cover three different kinds of behaviour which it may be difficult to distinguish: (*a*) the attempt to falsify a hypothesis *per se*, (*b*) implicit falsification of a hypothesis by an instance of a more general hypothesis which it entails, and (*c*) unintentional falsification of a hypothesis by a random instance. But these differences within the criterion are less important than those differences which clearly distinguish these kinds of behaviour from enumerative reasoning.

Wetherick's own criterion for eliminative reasoning (prediction that an instance will not conform to the unknown rule) is unsatisfactory. If a subject predicts that an instance, at variance with his own hypothesis, will not conform to the rule, then he presumably believes that his hypothesis is a likely candidate for the rule. Conversely, if he predicts the same instance will conform to the rule, he is presumably sceptical about his hypothesis. Hence, Wetherick's criterion implies that a subject can only be said to have eliminated a hypothesis if he had believed that this hypothesis was likely to be the rule, i.e. would not be eliminated. And this seems somewhat arbitrary. It makes the criterion of elimination depend on the degree of confidence which is felt about a hypothesis, and it would not count the checking of a dubious hypothesis as being eliminative reasoning at all.

In view of this criterion, it is not surprising that Wetherick finds much less evidence of elimination in his one group which started with an instance which was known to conform to the unknown rule.

In conclusion, the extent to which the phenomena reported by Wason are specific to the task, material and instructions used is, of course, an empirical matter which can only be settled by research. In this respect, Wetherick's modification of the original experiment is of interest.

BOOK REVIEWS

Sensory Communication. Edited by W. A. Rosenblith. M.I.T. Press and J. Wiley. 1961. Pp. 844. £6 8s.

This book contains 38 papers presented at a conference held at Endicott House, M.I.T., in July 1959. Though sponsored by the U.S. Air Force, this was the brain-child of Walter Rosenblith, who invited participants, organised the meetings, and has now edited the book. The idea was to bring together representatives from the various sciences interested in the working of the senses, in the hope that "Principles of Sensory Communication" would emerge. Forty-two took part, (18 from outside the U.S.A.) and these were distributed amongst different subjects as follows: Neurophysiology—21, Psychology—13, Zoology—3, Communication sciences—3, Physics—2. As one of the fortunate participants I can vouch for the excellence of the conference as an intellectual and social occasion, but what of the book?

In a brief review its scope is best illustrated by naming the authors and the subjects they are writing on, but it should not be assumed that the authors always run true to form: perhaps because of the mention of "principles of communication" in the title of the conference, some authors have let their hair down, and others gone to pieces. There are 5 chapters on general aspects of psychophysics and its relation to physiology. S. S. Stevens expounds the various forms of scaling that are possible and argues for the power law instead of the logarithmic law as the relation between physical and psychological magnitude: Ekman describes work in Sweden supporting the power law: Licklider describes the remarkable analgesic effect of loud sounds: Pollack discusses the contributions of information theory and signal-detection theory to psychophysics: Rushton discusses the relations between physical stimulus, nerve impulses, and sensation, and betrays a long concealed and treacherous leaning towards chemical as opposed to physical, explanations.

There are 7 chapters on hearing. H. Davis says that the peripheral apparatus does not itself resolve pitch, and Katsuki gives interesting results on the frequency sensitivity of units recorded from different levels in the cat's auditory pathways: Roeder and Treat describe the moth's supersonic bat-detecting organ: Woolsey describes the complex cortical representation of acoustic information: Neff describes auditory performance after ablations and destructive operations on the auditory system: Cherry describes a model in which pitch discrimination and localisation are performed by auto- and cross-correlation operations, and experiments supporting this model.

There are 8 chapters on vision. Ratliff gives a nice account of his admirable work with Hartline on lateral inhibition in Limulus: Goldsmith gives evidence on colour vision in the honey-bee: Bouman reviews work on quantum effects in vision: Boynton describes work on the enhanced brightness of short flashes, and on the variations of threshold near the onset of adapting lights: Lettvin, Maturana, Pitts, and McCulloch have a joint paper relating their electrophysiological findings to the histological structure of the frog's retina: Jung reviews work on recording from neurones in the striate cortex: Arden and Soderberg describe experiments showing what factors may influence the transmission of impulses through the lateral geniculate of rabbits: and MacKay describes various visual illusions.

Four chapters are on cutaneous sensation. Wall describes and compares records obtained from the second order cells in the dorsal horn of the spinal cord and in the nucleus gracilis: Mountcastle gives an account of units activated by cutaneous stimuli lying in the thalamus and somatic cortical areas I and II: Rosner is concerned with the inability to feel the second of two cutaneous stimuli delivered in quick succession, and tries to account for the failure physiologically: Geldard explores the possibilities of cutaneous stimulation as a means of communicating information to an individual.

Five chapters are devoted to taste and smell. Beidler discusses the mechanism of stimulation of the receptors, and de Vries describes some estimates of the number of molecules required to give rise to a sensation of smell: Landgren recorded from units in the thalamus and cortex that were activated by tactile, thermal, chemical, and noxious stimuli applied to the tongues of cats: Pfaffmann and collaborators studied the chemically activated units found in the rat medulla and thalamus: Zottermann recorded from the peripheral nerves in a wide variety of species, finding that there is a specific response to water in many species, and that the "sweet" fibres in the dog do not respond to saccharine.

Three papers are on factors which influence the response of central structures to sensory stimulation. Bremer discusses factors influencing evoked cortical potentials: Hernandez-Peon talks about the reticular control of sensory relays: Brazier and collaborators describe changes in evoked potentials produced by "conditioning."

Eight chapters remain to be described. One is an interesting experimental study by Buser and Imbert of units in the so-called "motor" cortex that are activated by sensory stimulation. Bullock and Fessard each grapple with the problem of how assemblies of neurones do what we know they do do, and Barlow discusses what relays in sensory pathways ought to do. Keidel and his colleagues argue that adaptation in a sensory system does not cause a loss of information, but is a resetting of the mechanism to improve discrimination in new circumstances. Rose and colleagues describe what may be a very important property of cortical neurones; with a beam of monoenergetic deuterons they damaged a layer at a certain depth in the cortex, and subsequently found regrowth of nerve terminals into the damaged area. Their animals were not, however, fully grown, so we still do not know whether mature cortical cells can continue to grow. Reichardt's paper describes how he was led to a certain mathematical model to explain Hassenstein's data on the optomotor responses of the beetle *Chlorophanus*. It seems to me especially interesting, first in giving a rather complete description of behaviour that is by no means simple, and second because the model is sufficiently detailed to provide definite physiological expectations if one could record from higher neurones in the beetle's visual pathways. Finally there is a chapter of comments from contributors, and one from the Editor himself: these are distillates of the lively discussions that occurred at the conference.

It will be seen that these 800 pages contain a lot of information, but it must be admitted that most of it has, or will, appear in the ordinary periodicals. Is this, then, another sop to our laziness in reading the scientific literature—a monument to the inefficiency of scientific communication? Or is it a useful collation of experimental results and theoretical ideas that achieve some unity, yet would not normally come the way even of an assiduous reader? As a participant I may be biased, but I think the latter would be a fairer judgement. It is not easy for people outside the field to understand the limitations and difficulties of neurophysiological investigations of the senses, and on the other hand it is easy for neurophysiologists to blind themselves to the greater problems of their subject in their efforts to solve the lesser. Binding in one volume contributions of people with such varied backgrounds may be a first step in bringing home, to contributors and others, that there is a single problem here. However it must be admitted that few solutions of the problem are to be found: the "principles" of sensory communication do not emerge, and it was wise to omit them from the title.

The book is nicely produced, but three warnings are needed. First, it is not a complete guide to the subject, nor would it be fair to expect this from a book which is the published proceedings of a conference. Second, the arrangement is rather haphazard; the book would have greater unity, and would be easier to find one's way about in, if the chapters were re-ordered and grouped. Thirdly, the price is high for a book that is unlikely to maintain its interest for more than five or ten years.

H. B. BARLOW.

Markov Learning Models for Multiperson Interactions. By Patrick Suppes and Richard C. Atkinson. London. Oxford University Press (for Stanford University Press.) 1960. Pp. xiv + 296. 66s.

This book is primarily a report on a large and ambitious four-year research programme. The authors have endeavoured "to apply a mathematical theory of behaviour to small-group experiments that closely resemble game situations." They report nine experiments; each experiment simulates a game in the sense of von Neumann and Morgenstern, but the players are given only partial information. During the course of an experiment the players acquire some of the missing information from their experience of repeated plays of the game. The authors develop and apply several related models, derived from the mathematical learning theory developed in the last decade by Estes and Burke, in their search for a quantitative description of this learning process.

The authors begin with a mathematical exposition of their basic theory. They have taken pains to keep the mathematics comparatively simple, and, as they suggest, little is required beyond a knowledge of discrete probability theory to understand their exposition. Particular developments of the theory are not described in this opening chapter, but are relegated to those later chapters in which the corresponding experimental data are presented. Even so, the mathematical complexity of this first chapter is greater than many psychologists would care to tackle. This is hardly a fault on the part of the authors; rather, any psychologist who wishes to penetrate into this field of work should be a mathematician as well. Such a conclusion is confirmed by the mathematical sophistication of the second chapter, in which the authors present certain statistical methods of analysis used in their research.

In their experiments the authors required their subjects to predict on each of a series of trials which of two lights would light up, and to indicate their prediction by pressing one of two keys. Two or three subjects, each with his own separate set of lights and keys, were run simultaneously. The "correct" answer on each trial was not necessarily the same for all the subjects, and was determined for each subject as a probabilistic function of the complete set of predictions on that trial and subsequent to those predictions being made. A pay-off matrix determined a priori the likelihood of each subject being "correct" conditional on each possible combination of responses. The subjects were informed of the "correctness" or otherwise of their predictions before beginning the subsequent trial, and this information was treated as a reinforcement in the sense of mathematical learning theory. Hence each trial of an experiment represented one play of a game, in which the two keys were the available strategies, the only possible outcomes were "win" and "lose," and the pay-off matrix was expressed in terms of probabilities of winning. In each experiment the subjects were kept in ignorance of some, or all, of the following factors: the pay-off matrix, their opponents' predictions, and the knowledge that they were even playing a game at all (*i.e.* the knowledge that their success depended upon another subject's predictions beside their own).

The data collected from these experiments consist of sequences of responses and reinforcements. For each trial the prediction made and the reinforcement received by each subject were recorded. The authors postulate that on each trial the response of a given subject is determined by the state of all-or-none conditioning of that subject. In their one-element θ-model there are two possible states of conditioning corresponding to the two available responses, and these states are directly observable. In between trials there may be changes in the conditioning states, and the model includes a specific hypothesis about such changes: after each trial the state of conditioning of a subject will change with probability θ if and only if the response reinforced is other than the one made on that trial. This requires that if the response made is reinforced it will always be repeated on the next trial—a prediction which is certainly not confirmed. But this model does provide accurate predictions of the asymptotic response probabilities, and these predictions are relevant elsewhere. θ is a constant peculiar to each subject. Most considerations require a value about 0·6, but the shape of the observed learning curve requires a value rather less than 0·05.

Since the states of conditioning of the one-element θ-model are directly observable, its conditioning axiom could have been tested directly. The authors have not done this explicitly, but they have pointed to this as the weakest aspect of the model and have introduced an appropriate alternative. The generalised conditioning model postulates four conditioning parameters for each subject specifying the probability of a change in the state of conditioning conditional on each response-reinforcement combination. These parameters are all estimated from the data and the necessary and sufficient condition for the matrix (2.6.9.) to fit the data is that the changes in the states of conditioning of the two subjects shall be statistically independent. At asymptote this is indeed so, as is evidenced by the statistics quoted in Table 12.2. This feature of statistical independence is a basic ingredient in every model discussed in this book. But the authors fail to point this out, nor do they give an adequate discussion of the statistical independence inherent in their models. Their analysis of their data would have been more rational if this test had come first instead of last in order.

The generalised conditioning model has one other necessary condition: the Markov process of response random variables must be stationary and of the first order. On examination of their data the authors found that the Markov chain was stationary only at asymptote, and that some tests of second order against first order processes were significant.

The excellent fit obtained with the transition matrix of the generalised conditioning model does not indicate a superiority over the θ-model, but that it places weaker requirements on the data. The point of interest is now the conditioning parameters. These are heavily influenced by the reinforcement schedule in operation, and are different for the two players. Very puzzling is the fact that this difference still holds in Chapter 5, where the game situation is symmetric between the two players. The authors examine the changes occurring in the conditioning parameters with successive trials for an experiment in which monetary pay-offs were used with a non-contingent reinforcement schedule. They find that the making of a particular response has a "reinforcing" value comparable to the knowledge of which response was "correct," and that as the monetary reward increases, so the relative effect of this knowledge of results decreases. These results appear paradoxical only if we think in terms of a Pavlovian type of conditioning. The concept of conditioning applied here is nothing more than an all-or-none relation between the response and reinforcement on one trial and the response on the next, and has nothing in common with that developed by Pavlov but the name.

Another way out of the difficulties of the one-element model is provided by postulating more than one stimulus element. On each trial the experimenter presents his subjects with a stimulus situation which, the authors suppose, contains several stimulus elements; each subject samples a subset of these. This is very much a mathematical generalisation. The authors make no attempt to identify their stimulus elements with particular aspects of their experimental situation. It follows that in the multi-element θ-model the states of conditioning are no longer directly observable, and the Markov process is of infinite order, allowing some sequential effects to be predicted. It also happens that the mathematics of this model is exceedingly complicated and tedious, and the authors have derived detailed predictions in only two particular cases.

The asymptotic response probability predictions are substantially the same as for the single-element model; hence these predictions are important even though the single-element model fails dramatically in other respects. The multi-element model does not require the Markov process to be stationary to the first order except at asymptote, and it predicts second- and higher-order effects, although these will not necessarily be significant in any realisable experiment. The statistical tests of the multi-element model give variable results. The predicted transition matrix fits a non-contingent reinforcement schedule well, but shows a highly significant departure in a zero-sum game experiment. The predicted learning curve approaches asymptote much too quickly, and the variance in the observed asymptotic response probabilities is rather too large. The accuracy of the sequential predictions falls off for orders greater than the first.

In their book the authors examine briefly several other models, notably a Linear Model and a Hypothesis Model, but find, in general, that these models do not fit the data so well as the stimulus-sampling models. Of these latter models the one-element θ-model is the most widely tested, and this is applied to game situations in which there is communication between the players, others in which the pay-off matrix is displayed, and finally to an experiment on utility functions.

The three stimulus-sampling models, which have been examined closely in this review, are the most successful, but yet none of them can be called satisfactory. Each of them gives some successful predictions and others which give highly significant departures. The most important result is the independence of the first order transitions in the response-reinforcement random variables. This demonstrates that the concept of interaction used by the authors—that the mutual influence of the players acts solely through the reinforcements which each receives—is appropriate in this context. It suggests that no relevant variable has been omitted in constructing the models.

Any estimation of this work must hinge on the question whether the theories described here are ever likely, with modifications, to be able to account accurately for the learning processes studied, or whether they contain fundamentally inappropriate postulates. It is not possible to give an objective answer to this question, since the book contains many hopeful, yet unexplored, suggestions and far too few successful results. Any answer which the reviewer or anyone else might give would reflect far more his interest in and involvement with mathematical learning theory than it would the contents of this book. Suffice it to say that the sceptics are justified in remaining sceptical.

D. R. J. Laming.

Ergonomics. Vol. 5. *No.* 1. *January*, 1962. London. Taylor and Francis. Pp. 334. £2 10s.

This special issue is devoted to the Proceedings of the First Congress of the International Ergonomics Association held at Stockholm in August, 1961.

O. L. Zangwill.

INDEX OF AUTHORS

Anthony, W. S., with Holding, D. H., Lion, J. S., and Sluckin, W. 77
Berger, R. J., with Olley, P., and Oswald, I. 183
Bitterman, M. E., with Gonzalez, R. C. 109
Blancheteau, M., with Fraisse, P. 52
Brebner, J., with Gordon, I. 113
Burgh, P., with Grindley, G. C. 89
Carr, A. E., with Sutherland, N. S. 140
Carr, A. E., with Sutherland, N. S., and Mackintosh, J. A. 129
Conrad, R. 71
Corcoran, D. W. J. 178
Dale, H. C. A. 49
Davis, R. 157
Devine, J. H., with Gregory, A. H. 122
Dodwell, P. C. 65
Fraisse, P., with Blancheteau, M. 52
Gibbins, K., with Howarth, C. I. 167
Gonzalez, R. C., with Bitterman, M. E. 109
Gordon, I., with Brebner, J. 113
Gregory, A. H., with Devine, J. H. 122
Grindley, G. C., with Burgh, P. 89
Herberg, L. J. 8
Heron, A. 94
Holding, D. H., with Anthony, W. S., Lion, J. S., and Sluckin, W. 77
Howarth, C. I., with Gibbins, K. 167
Hurwitz, H. M. B. 1
Kinsbourne, M., with Warrington, E. K. 223 and 235
Laming, D. R. J. 38
Lion, J. S., with Anthony, W. S., Holding, D. H., and Sluckin, W. 77
Mackay, C. K., with Symons, J. R. 104
Mackintosh, J. 15
Mackintosh, J. A., with Sutherland, N. S., and Carr, A. E. 129

Mowbray, G. H. 117
Olley, P., with Berger, R. J., and Oswald, I. 183
Oswald, I., with Berger, R. J., and Olley, P. 183
Senders, J. W., with Senders, V. L. 176
Senders, V. L., with Senders, J. W. 176
Sluckin, W., with Anthony, W. S., Holding, D. H, and Lion, J. S. 77
Sutherland, N. S., with Carr, A. E. 140
Sutherland, N. S., with Carr, A. E. and Mackintosh, J. A. 129
Symons, J. R., with Mackay, C. K. 104
Treisman, A. 23
Warrington, E. K., with Kinsbourne, M. 223 and 235
Wason, P. C. 250
Wetherick, N. E. 246
Wilson, J. 119
Young, J. Z. 193 and 206